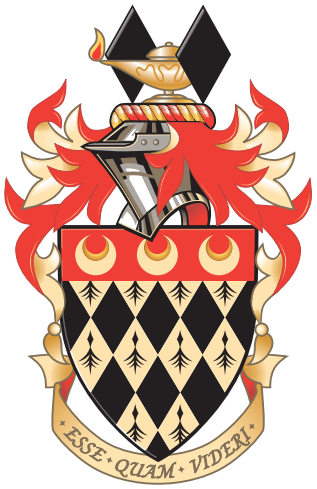
Searches for Non-Resonant New Physics in Dielectron Events with ATLAS

Graham Savage

Department of Physics

Royal Holloway, University of London



A thesis submitted to the University of London for the Degree of Doctor of Philosophy

June 2, 2016

DECLARATION

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the document.

Graham Savage

Abstract

Two searches for non-resonant new physics phenomena are presented using the data collected with ATLAS in proton-proton collisions at the Large Hadron Collider. The 2012 $\sqrt{s} = 8$ TeV dataset is used to search for large extra spatial dimensions and the 2015 $\sqrt{s} = 13$ TeV dataset is used to search for a four fermion contact interaction. These datasets correspond to 20.3 fb^{-1} and 3.2 fb^{-1}, respectively. The searches are performed using events with dielectron final states, and combined with dimuon events for a dilepton result. Due to an absence of signal, 95% credibility levels are set on the parameters of interest using a Bayesian interpretation. Lower limits are set on the string scale M_{S} for large extra spatial dimensions between 3.0 TeV and 5.0 TeV. Lower limits are set on the contact interaction scale Λ between 14.4 TeV and 25.2 TeV.

Acknowledgements

I would like to thank many people for their help and guidance throughout my PhD. Firstly, my supervisor Tracey Berry who has provided me with support and the opportunities to get the most out of my PhD.

I extend my gratitude to the Science and Technologies Facilities Council, without whose financial support my research would not have been possible.

I am indebted to Royal Holloway academics; Glen Cowan, Veronique Boisvert and Pedro Teixeira-Dias for four years of advice and discussion of my research. I would also like to thank Stephan West for his help on exotic particle physics theory.

My thanks go to the entire run 1 exotics non-resonant group, especially Liam Duguid, Marc Cano Bret and Tülin Varol. I could not have asked for a better group of people to work with and learn from. A special thanks goes to Dan Hayden, who patiently answered all the questions that a new ATLAS student has and guided me through both the analyses on which I have worked.

Thanks to Claire, Josh and Ian from the Royal Holloway office for the fun, friendship and general discussions along the way. This thanks is further extended to Russell, who has taken much time to spell out theoretical concepts for my benefit, and to Ben for keeping me company in all our CERN based adventures.

Finally I would like to thank my family for their continued encouragement, without this the 4 years would have been much more of a challenge.

Contents

List of Figures

List of Tables

Preface

High energy particle physics experiments are incredibly complex, requiring world wide collaboration in funding and operation. ATLAS alone has a team of over 3000 people. While this thesis details the work undertaken by myself, it would be naïve to think any of this could be carried out without a combined effort. Here a brief description of what is included in this thesis is given, highlighting my involvement in the two analyses and where others have provided input.

This thesis presents two searches for non-resonant new physics (ADD large extra dimensions and four-fermion contact interactions (CI)) using the ATLAS detector. These searches are conducted with the 2012 $\sqrt{s} = 8$ TeV and 2015 $\sqrt{s} = 13$ TeV datasets, respectively. I contributed majorly to the dielectron channel in both searches. The dielectron channel is complementary to the dimuon channel and can be combined for a dilepton result. Where work is presented for a dilepton result, the dimuon inputs were provided by either Marc Cano Bret or Tülin Varol.

For the 2012 result I focused mainly on the ADD search [1]. My main role was implementing and running the event selection for data-background comparisons, yield tables and cross-checks, then the optimisation, search and limit setting phase in the dielectron and dilepton channels. The combination with the muon channel in this search relied upon the inputs provided by Tülin Varol [2] and Marc Cano Bret [3]. For the 2015 result I worked on many aspects of the dielectron channel and on the CI search in general [4]. I performed much of the event selection checks on signal efficiency, and was a main analyser in dielectron data-background comparisons. I provided the nominal and systematic uncertainty variations of the dielectron event selection for the working group, and created the dielectron expected event yield tables. I led the CI dielectron and combined limit setting phase, including the maintenance of the limit setting infrastructure. Once again the combination with muon channel relied upon input from Marc (as documented in [4]).

Not included in this thesis is the work undertaken to achieve ATLAS authorship status. Working in the electron/photon group my main role was to migrate an existing trigger inefficiency tool into a new software framework. This also involved validating samples and checks of variable distributions for the new data format with quick iteration with the

central production experts of any issues.

In this thesis Chapters 1–3 focus on the theory and experimental methods behind the searches. Chapter 4 describes the data samples used in the searches. Chapters 5–9 detail my contributions to the two non-resonant searches [1, 4]. At the time of submission, the evolution of the results in the conference note [4] are still in the paper review process. This thesis contains all the updates for the paper results which are not present in the conference note. A brief outline of each chapter is given below.

Chapter 1: Theoretical Motivation

This chapter provides an overview of the Standard Model of particle physics, highlighting the aspects from which backgrounds to these searches arise. The theoretical motivations for ADD gravitons and CI are given, and a description of how these models would be observable at the Large Hadron Collier if present in nature. This chapter concludes with a discussion of the current experimental limits on these models.

Chapter 2: Experimental Setup

In this chapter a description of the Large Hadron Collider and how it provides the ATLAS experiment with high energy proton-proton collisions is overviewed. Details of the ATLAS experiment are given, concentrating on the detection of electrons which are the focus of this thesis. A brief review of the trigger system and the performance of ATLAS is given.

Chapter 3: Electron Reconstruction

This chapter reviews how electrons are reconstructed from the measured ATLAS detector signals, and how the novel construct of the Inner Detector and Electromagnetic Calorimeter can separate signal from background. Also included is a description of the criteria used in building electron candidates with different levels of purity.

Chapter 4: Background and Signal Estimation

Within this chapter the methods of estimating the backgrounds and signals used in these searches is described. The validation of the CI signal re-weighting is discussed here.

Chapter 5: Event Selection

This chapter details the event selection used in each analysis. In the $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV analyses the data-background comparisons and the expected yields are given. In the $\sqrt{s} = 13$ TeV analysis the effect of using track-vertex matching parameters is studied. An investigation into the choice of electron working points is presented.

Chapter 6: Systematic Uncertainties

The sources of theoretical and experimental uncertainties are described in this chapter. In the $\sqrt{s} = 8$ TeV analysis a study of including an additional background sample is performed, the results of which are used in the statistical analysis. The effect on the expected event selection depending on systematic variations is investigated.

Chapter 7: Statistical Analysis

This chapter introduces the statistical methods used to quantify a discovery and set limits on the parameters of interest. A motivation for the choice of prior is given. The method of using parameterisations to model signal is discussed. The investigation into an optimised search range used in ADD analysis is explained.

Chapter 8: Results

This chapter describes the results from the statistical analysis. Presented are the results from the $\sqrt{s} = 8$ TeV dielectron and dilepton ADD large extra dimensions search, and the $\sqrt{s} = 13$ TeV dielectron and dilepton CI search.

Chapter 9: Conclusion

Finally this chapter concludes the findings from these searches.

"We all agree that your theory is crazy, but is it crazy enough?"

– Niels Bohr

1

Theoretical Motivation

The Standard Model (SM) of particle physics is a predictive theory that has been experimentally verified up to the TeV energy scale obtainable at modern day particle colliders. Fortunately for the current generation of particle physicists, the SM is known to have limitations, and much work has gone into searching for possible solutions. This chapter begins with a brief overview of the SM including a description of the fundamental particles and forces, and how these form the backgrounds to the searches presented. Two Beyond the Standard Model (BSM) theories are discussed which provide possible extensions to the SM. Finally the most recent tests of these models are summarised.

1.1 The Standard Model

The SM is a quantum field theory (QFT) which represent particles of nature as excitations of fields. A theory invariant under the gauge transformations of $SU(3)_C \times SU(2)_L \times U(1)_Y$ leads to a description of all fundamental forces observed so far in nature except gravity. The SM particles are either integer spin bosons or half-integer spin fermions. The bosons consist of spin-1 gauge bosons which mediate the three predicted forces; electromagnetic, weak and strong interactions, and the spin-0 Higgs boson. The spin-$\frac{1}{2}$ fermions make up matter and can be further differentiated into leptons and quarks each containing three generations or flavours.

1.1.1 Fundamental Particles

Leptons

The SM contains six leptons (ℓ), each with a corresponding antilepton ($\bar{\ell}$). The three flavours of charged leptons are; electron (e^-), muon (μ^-) and tau (τ^-). Each charged lepton has a related neutral lepton; the electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_τ). Charged leptons interact through the electromagnetic and weak forces, while neutrinos through the weak force only. The properties of leptons are summarised in Table 1.1.

Charge (Q)	Generation		
	1	2	3
-1	electron e $m = 0.511$ MeV	muon μ $m = 105.7$ MeV	tau τ $m = 1.777$ GeV
0	electron neutrino ν_e $m \approx 0$ eV	muon neutrino ν_μ $m \approx 0$ eV	tau neutrino ν_τ $m \approx 0$ eV

Table 1.1: Summary of the Standard Model leptons [5].

Quarks

The SM consists of six quarks (q); up (u), down (d), charm (c), strange (s), top (t) and bottom (b), each with a corresponding oppositely charged antiquark ($\bar{\text{q}}$). Quarks interact through the strong, weak and electromagnetic forces. Quarks have an inherent property called "colour" which the strong interaction couples to (discussed in Section 1.1.2). This allows quarks to form hadrons, grouped into baryons (qqq) and mesons ($\text{q}\bar{\text{q}}$), as long as the sum of the colours renders the hadron "colourless". Protons are the only stable baryons containing the quark configuration uud. Many other configurations of quarks form different baryons and mesons but none are stable. The properties of quarks are summarised in Table 1.2.

Gauge Bosons

The vector gauge bosons in the SM are the photon (γ) mediating the electromagnetic force, the Z and $W^{\pm}$ bosons mediating the weak force and the gluon (g) mediating the strong force. The W boson is charged and is therefore the only boson distinguishable from its anti-particle. The properties of the gauge bosons are summarised in Table 1.3.

1.1.2 Fundamental Forces

Each interaction and particle in the SM is represented as a field, and the entire system described by a single Lagrangian. A gauge transformation is the changing from one field configuration to another. It is a requirement that a gauge transformation leave the Lagrangian invariant for the theory to be renormalisable and predictive of observables. Under a given gauge transformation a new gauge boson field must be introduced to leave the Lagrangian invariant. These can be interpreted as the mediating particles of the interaction. The introduction of a gauge boson mass term is not invariant, and as such forbidden in the Lagrangian under the associated gauge transformation. Physically these transformations require all gauge bosons to be massless.

Quantum Electrodynamics

Quantum Electrodynamics (QED) is the QFT describing the electromagnetic interaction of charged particles mediated by the emission and absorption of massless photons. Formally, QED is invariant under gauge transformations of the $U(1)$ group with the introduction of a photon field, and leads to the conservation of electric charge. This interaction acts over infinite range coupling to electric charge. There is no self coupling of this interaction as the photon is not electrically charged. QED is of particular importance in the searches presented as it contributes to the dominant Drell–Yan (DY) background process ($q\bar{q} \rightarrow \gamma^* \rightarrow \ell\bar{\ell}$).

Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the QFT describing the strong nuclear force binding quarks together in hadrons. Gauge invariance in QCD is described by the $SU(3)_C$ group and leads to the introduction of 8 massless gluons acting as mediators of the strong force

Charge (Q)	Generation		
	1	2	3
+2/3	up u $m \approx 2.3$ MeV	charm c $m \approx 1.275$ GeV	top t $m \approx 173$ GeV
-1/3	down d $m \approx 4.8$ MeV	strange s $m \approx 95$ MeV	bottom b $m \approx 4.18$ GeV

Table 1.2: Summary of the Standard Model quarks [5].

Force	Charge (Q)	Boson
Electromagnetic	0	photon γ $m = 0$
Weak	0	Z boson Z $m = 91.2$ GeV
	$\pm$ 1	W boson W $m = 80.4$ GeV
Strong	0	gluon g $m = 0$

Table 1.3: Summary of the Standard Model gauge bosons [5].

and the conserved quantity of colour charge. As gluons carry colour charge they can self interact. Leptons do not carry colour charge and as such do not experience the strong interaction. Unlike the electromagnetic interaction the potential increases as distance grows between two colour connected quarks — a phenomenon know as confinement. If the energy between a quark pair increases sufficiently, two new quarks are created, making two sets of colour connected quarks. This process continues until the quarks do not have sufficient kinetic energy to separate and produce new pairs, instead the quarks hadronise. The result of this hadronisation is a collimated collection of hadrons known as a jet. This illustrates how only colour neutral states are seen in nature and not free quarks. Jets make up the reducible fake electron background in this search. In contrast, at very short distances quarks behave as free particles in a phenomenon known as asymptotic freedom.

Weak Interaction

The weak interaction is unique in that it involves all quarks and leptons. The two types of weak interaction are the charged current interaction, mediated by the $W^{\pm}$ bosons and the neutral current interaction, mediated by the Z boson. Charged current interactions are the only mechanism allowing quarks to change flavour at tree level and are responsible for radioactive decay. Due to the massive nature of these mediators, the weak force has the shortest range of all SM forces, 10^{-18} m. Neutral current interactions through the Z boson are of particular importance to these searches as they contribute to the DY background ($q\bar{q} \rightarrow Z \rightarrow \ell\bar{\ell}$). W and Z mediators can self couple, which comprises part of the non-negligible diboson background process of ZZ, WZ and WW decaying leptonically. The experimental non-zero W and Z mass measurements showed the weak theory had to be spontaneously broken as the massless nature of the gauge boson field is otherwise essential to preserve gauge invariance.

Electroweak Unification and Spontaneous Symmetry Breaking

The electromagnetic and weak interactions were successfully unified by Glashow [6], Salam [7], and Weinberg [8] in the 1960s. The electroweak interaction is based on the gauge transformation of the $SU(2)_L \times U(1)_Y$ groups above the electroweak (EW) unification scale. As a consequence, four gauge fields are introduced with the conserved quantity

hypercharge. It is the mixing of these fields that result in the observable $W^{\pm}$, Z and γ gauge bosons.

As it is impossible to explicitly introduce a mass term into the Lagrangian without loosing gauge invariance, another mechanism was theorised — spontaneous symmetry breaking through the Englert–Brout–Higgs–Guralnik–Hagen–Kibble mechanism [9–11]. A new complex scalar doublet field is introduced, the Higgs field, in which particles interact to acquire mass. The potential for the new field, can be written as:

$$V(\phi) = \mu^2|\phi|^2 + \lambda|\phi|^4 \ . \tag{1.1}$$

If $\mu^2 > 0$ the potential is parabolic with a minimum at $\phi = 0$. If $\mu^2 < 0$, the minima is instead found at $\phi \neq 0$. This leads to a potential which is not symmetric from the ground state perspective. The Lagrangian of the system is invariant under gauge transformations whereas the ground state is not, so the symmetry is said to have been spontaneously broken. It is from the Higgs mechanism that fermions and weak gauge bosons acquire mass. The Higgs Boson is an excitation of this field and was the final piece of the SM discovered in 2012 [12, 13].

1.1.3 Parton Distribution Functions

In high energy proton-proton (pp) collisions, the relativistic hadrons appear as a collection of point-like constituents with a wide-spread momentum distribution. The momentum distribution functions of the partons are called Parton Distribution Functions (PDFs) which represent the probability density of finding a parton carrying a momentum fraction, x, at a given squared momentum transfer, Q^2 ($f(x, Q^2)$). In Deep Inelastic Scattering (DIS) experiments, the nucleon was found to have a substructure of three valence quarks imbedded in a sea of gluons and virtual quark-antiquark pairs. Collectively, these constituents are named partons. At low Q^2, the three valence quarks become dominant carrying the largest share of momentum. At high Q^2, there are more sea quarks each carrying a low momentum fraction. DIS experiments showed that only half of the nucleon momentum is carried by quarks, the remainder carried by gluons. The PDFs are described in the QCD

evolution equations, known as the DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) equations. The absolute values of the PDFs cannot be calculated in perturbative QCD, rather the x dependence of the parton distributions at a given Q^2 needs to be determined from experiment. To eliminate infrared and ultraviolet divergences in perturbative QCD calculations, factorisation and renormalisation scales are introduced, respectively. These make calculations possible between a given energy regime, and need to be chosen prior to calculation.

In order to produce high invariant mass particles, the colliding partons must carry a significant fraction of the protons momentum. As the amount of data is limited at high x from experiment, the error on the PDFs are large in this kinematic region. As such, these are the largest uncertainties in the presented high invariant mass searches. At higher centre-of-mass energies, the x range shifts to lower values for the production of particles at a given mass when compared with lower centre-of-mass energies. As the PDFs are better constrained for these values the uncertainty decreases for higher energies and the corresponding searches benefit from the reduced systematic uncertainty.

1.2 Beyond the Standard Model

Whilst the SM has proven to be a useful tool in understanding nature, it is not without limitations. One of the main physics programmes at ATLAS (A Toroidal Lhc ApparatuS) is to search for phenomena predicted by BSM theories. With world record energies being reached in pp collisions at the Large Hadron Collider (LHC), BSM theories have faced rigorous testing, often using the dilepton (two electrons or muons in the final state) invariant mass ($m_{\ell\ell}$) spectrum as one of the cleanest observables in which to search for deviations from the SM. Often these deviations would come in the form of narrow resonances, such as a massive new gauge boson at the TeV scale. This thesis concentrates on two searches for non-resonant processes (large extra dimensions (LED) at $\sqrt{s} = 8$ TeV (where $\sqrt{s}$ is the centre-of-mass energy) and four-fermion contact interactions at $\sqrt{s} = 13$ TeV) in the high invariant mass dielectron (m_{ee}) spectrum. If present in nature these processes would be observable in data as a broad excess over the SM prediction.

1.2.1 Large Extra Dimensions

One of the major outstanding questions in particle physics is the vast difference between the Planck scale ($\sim 10^{16}$ TeV) — where the effects of quantum gravity are expected to become dominant, and the EW scale ($\sim$1 TeV) — where the electromagnetic and weak forces are unified. This phenomenon is known as the "hierarchy problem". Arkani-Hamed, Dimopoulos, and Dvali (ADD) put forward a model in 1988 providing a solution to the hierarchy problem with the introduction of n additional spatial dimensions, each compactified with radius R on an n-dimensional torus [14]. While SM particles are confined to the 3+1 space-time dimensions known as a brane, gravitons propagate through all dimensions of the universe (the bulk). It is the dilution of gravity as it passes through these extra dimensions that explains its apparent weakness compared to the three SM forces. A fundamental Planck mass (M_{D}) is introduced which can be related to the observed Plank scale (M_{Pl}) through the presence of these extra dimensions by;

$$M_{\mathrm{Pl}}^2 = M_{\mathrm{D}}^{n+2} R^n. \tag{1.2}$$

With a large enough radius for a given n, the fundamental Planck scale can be reduced to the order of the EW scale. These are described as large extra dimensions when compared to the size of proton, where $n = 1$ would require $R \sim 10^{11}$ m (and has been excluded), and for $n = 2$ the size decreases to $R \sim 100$ µm – 1 mm.

As these are compact extra dimensions, the graviton field has to have the same amplitude when entering and leaving the additional dimension. As such gravitons can only take on discrete momenta in these extra dimensions, which appear in the 4-d effective theory as a set of graviton mass modes known as a Kaluza–Klein (KK) tower [15]. The splitting of these mass states is proportional to $1/R$, thus due to the large nature of these dimensions the splitting is fine resulting in a continuous spectrum of KK states. Summing over these KK modes predicts an observable effect: a broad increase over the DY spectrum at high dilepton invariant masses. This sum leads to an integral that has to be regulated by an ultraviolet cutoff.

The ADD model is a low energy effective theory which is valid to the string scale, M_{S}, characterising the onset of quantum gravity. It is standard to equate the cutoff value with

the scale M_S, which is related to the fundamental Planck scale M_D by [16];

$$M_S = 2\sqrt{\pi}[\Gamma(n/2)]^{1/(n+2)} M_D. \tag{1.3}$$

Graviton exchange can result from two processes, a $q\bar{q}$ initiated process in analogy with the DY process, or by a gluon-gluon (gg) collision where there is no SM equivalent. Leading order processes for virtual graviton production are shown in Figure 1.1.

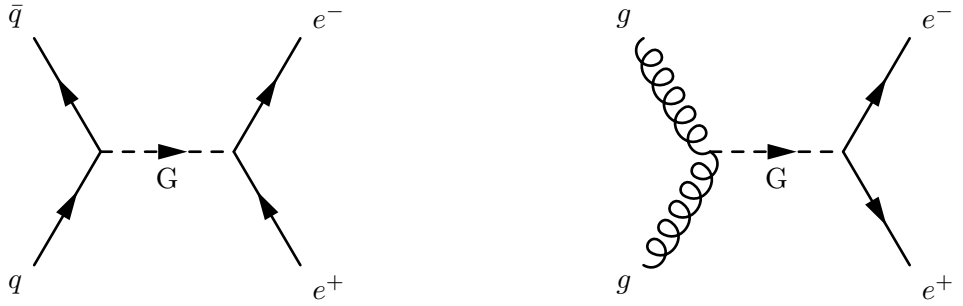


Figure 1.1: Leading-order production mechanisms for virtual Gravitons decaying to a dielectron final state.

The total cross-section of the SM DY process plus the ADD new physics contribution is given in Equation 1.4;

$$\sigma_{\text{tot}}(m_{\ell\ell}) = \sigma_{\text{DY}}(m_{\ell\ell}) + \mathcal{F}\frac{F_I}{M_S^4} + \mathcal{F}^2\frac{F_G}{M_S^8}\ , \tag{1.4}$$

where the first term is the pure SM DY process, and F_I and F_G are functions of the cross-section denoting the interference and pure graviton contributions respectively. The strength of the interaction is parameterised by $\mathcal{F}/M_S^4$, where $\mathcal{F}$ is a formalism dependent parameter. In this analysis, three formalisms are used, Giudice, Rattazzi, and Wells (GRW) [17], Hewett [18], and Han, Lykken, and Zhang (HLZ) [19]. The values of $\mathcal{F}$ for each formalism are:

$$
\begin{aligned}
\mathcal{F} &= 1 && \text{(GRW)},\\
\mathcal{F} &= \begin{cases} \log\left(\frac{M_S^2}{m_{\ell\ell}^2}\right) & n = 2 \\ \frac{2}{n-2} & n > 2 \end{cases} && \text{(HLZ)},\\
\mathcal{F} &= \frac{2\lambda}{\pi} = \frac{\pm 2}{\pi} && \text{(Hewett)}.
\end{aligned} \tag{1.5}
$$

While the HLZ formalism is dependent on the number of extra dimensions, the GRW and Hewett conventions do not vary with n. $\mathcal{F}$ is always positive when considering GRW or HLZ cases, however in the Hewett formalism the Graviton can interfere destructively as well as constructively with the SM DY process, this is chosen by the sign of the unity quantity λ, which is positive for constructive and negative for destructive interference. The interference effects between the DY and virtual graviton processes are small at the LHC due to dilepton production by virtual KK gravitons being predominantly gluon-induced rather than quark-induced (discussed further in Section 7.1.1), as such only the constructive case is considered.

This model decays to electrons and muons with an equal branching fraction. The decay to photons has twice the rate to a given lepton. The expected high mass excess above the SM DY from this model can be seen versus the dielectron truth mass in Figure 1.2 for a set of generated M_S samples under the GRW formalism. As M_S increases the new physics cross-section converges to the SM DY value.

1.2.2 Contact Interactions

At energy scales unachievable at the LHC, new phenomena not directly observable may become detectable as deviations from the SM predictions. The experimental sensitivity of new physics at scales beyond the available centre-of-mass energy can be parameterised using a four-fermion contact interaction (CI) theory. This approach was successfully formulated by Fermi to describe beta decay long before the discovery of the W boson. The additional CI terms could arise from either fermion substructures or a new massive particle.

In the SM, quarks and leptons are point-like particles in nature. The SM fails to explain the variety of observed quark and lepton flavours, their difference in masses, and

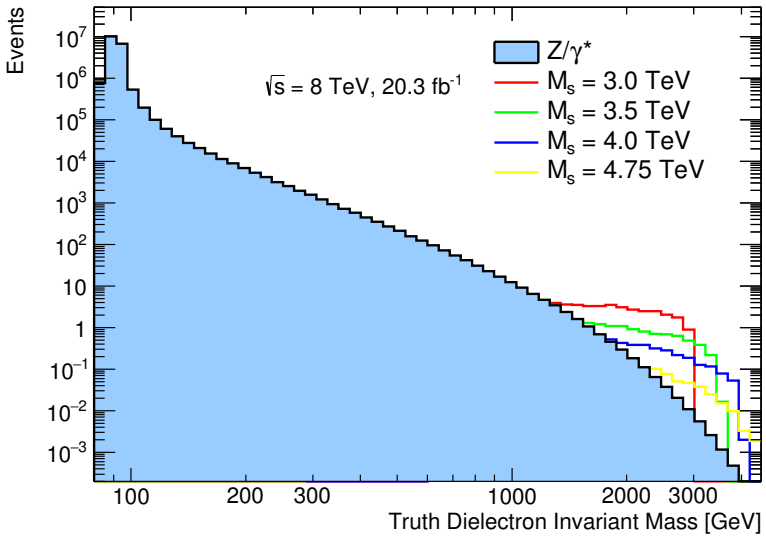


Figure 1.2: Generator level ADD dielectron invariant mass spectrum.

why these particles fall into a pattern of three generations. One explanation could be quarks and leptons are in fact made of more fundamental constituents, called preons [20], bound tightly together via a new gauge interaction. At collision energies exceeding the binding energy of this new interaction, multiple production processes would dominate over SM processes. If the energy scale is unachievable for direct observation, deviations in the dilepton mass tail could still be visible.

The effective Lagrangian describing a new interaction between two incoming quarks and a two lepton final state is [21];

$$\begin{aligned}\mathcal{L} = \frac{g^2}{\Lambda^2}[\ &\eta_{\mathrm{LL}}\ (\bar{q}_{\mathrm{L}}\gamma_\mu q_{\mathrm{L}})\ (\bar{\ell}_{\mathrm{L}}\gamma^\mu \ell_{\mathrm{L}})\ +\ \eta_{\mathrm{RR}}\ (\bar{q}_{\mathrm{R}}\gamma_\mu q_{\mathrm{R}})\ (\bar{\ell}_{\mathrm{R}}\gamma^\mu \ell_{\mathrm{R}}) \\ &+\ \eta_{\mathrm{LR}}\ (\bar{q}_{\mathrm{L}}\gamma_\mu q_{\mathrm{L}})\ (\bar{\ell}_{\mathrm{R}}\gamma^\mu \ell_{\mathrm{R}})\ +\ \eta_{\mathrm{RL}}\ (\bar{q}_{\mathrm{R}}\gamma_\mu q_{\mathrm{R}})\ (\bar{\ell}_{\mathrm{L}}\gamma^\mu \ell_{\mathrm{L}})]\ ,\end{aligned} \tag{1.6}$$

where g is a coupling constant chosen such that $g^2/4\pi = 1$, Λ is the energy scale of the contact interaction, and $q_{\mathrm{L,R}}$ and $\ell_{\mathrm{L,R}}$ are left-handed and right-handed quark and lepton fields, respectively. The chiral structure of the interaction is defined by η_{ij}, where i and j are L or R (left or right). The four possible chiral structures are investigated here separately, for example the left-left model is obtained by setting $\eta_{\mathrm{LL}} = \pm 1$ and

$\eta_{\mathrm{LR}} = \eta_{\mathrm{RL}} = \eta_{\mathrm{RR}} = 0$. The sign of η_{ij} determines whether the interference is constructive ($\eta_{ij} = -1$) or destructive ($\eta_{ij} = +1$). The cross-section for the process $q\bar{q} \to \ell\bar{\ell}$ including the contribution of these contact interaction models is then;

$$\sigma_{\mathrm{tot}}(m_{\ell\ell}) = \sigma_{\mathrm{DY}}(m_{\ell\ell}) - \eta_{ij}\frac{F_{\mathrm{I}}}{\Lambda^2} + \frac{F_{\mathrm{C}}}{\Lambda^4} \ , \tag{1.7}$$

where the first term is the SM DY process, the second term corresponds to the interference between the DY and CI processes, and the third term describes the pure CI process. The CI terms include parameters F_{I} and F_{C} respectively; these are functions of the differential cross-section with respect to $\mathrm{m}_{\ell\ell}$, but have no dependence on Λ. The leading order production mechanism for CI and a comparison to the SM DY process can be found in Figure 1.3.

Figure 1.3: Leading-order Feynman diagrams for Drell–Yan and additional Contact Interaction process with scale Λ decaying to a dielectron final state.

The observable excess at high dielectron invariant mass for a few CI models can be seen in Figure 1.4, as the scale of Λ increases the cross-section approaches the DY value.

1.2.3 Previous Searches

Large Extra Dimensions

Previous searches for ADD LED have been performed at many different collider experiments. The strongest observed 95% credibility level (C.L.) lower limits set at each collider are described below for the benchmark GRW formalism and are summarised in Table 1.4.

At electron-positron collider experiments [22–26], the angular distribution of the final state particles was used as a discriminating variable. Due to the asymmetry of the Z bosons coupling to left and right handed fermions, the angular distribution of the particles

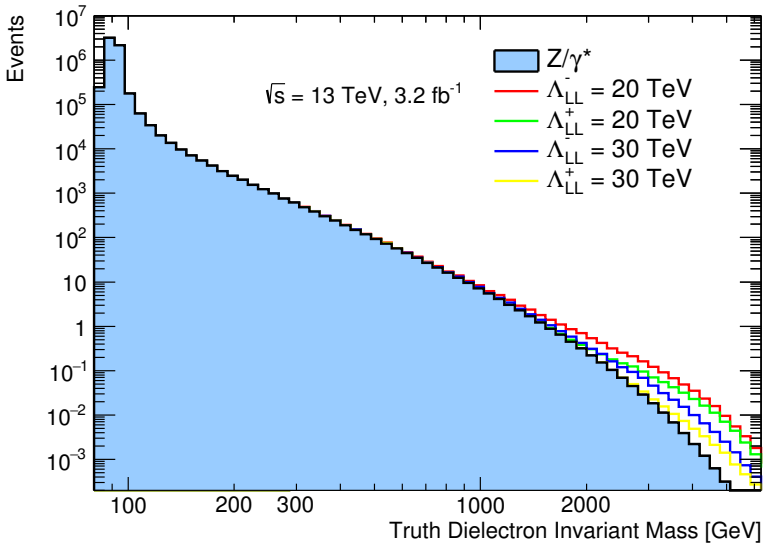


Figure 1.4: Generator level CI dielectron invariant mass spectrum.

in the final state can be used to discriminate signal from background. It can also be used to determine the spin of the intermediate particle — important when searching for a spin-2 graviton. The strongest lower limit on the mass scale was set in diphoton events at $M_S > 880$ GeV [26].

At the electron-proton collider experiment HERA, squared momentum transfer (Q^2) was used to determine a lower limit at $M_S > 720$ GeV [27].

At the Tevatron, searches have been conducted using the invariant mass and angular distributions of dielectron, dimuon and diphoton events [28,29]. The strongest lower limits arise from a two electromagnetic object search where the lower limit was set on the mass scale at $M_S > 1.62$ TeV [28].

The most recent searches for ADD LED have been performed at the LHC. ATLAS [30, 31] and CMS [32–34] both used the dielectron, dimuon and diphoton invariant mass distributions to search for new physics. ATLAS set the strongest lower limits using the $\sqrt{s} = 7$ TeV (5 fb^{-1}) dataset at $M_S > 3.22$ TeV in the combined dielectron, dimuon and diphoton channel. The lower limits from this dataset in the dielectron channel are $M_S > 2.73$ TeV, in the dimuon $M_S > 2.83$ TeV and in the combined dilepton channel $M_S > 3.00$ TeV. CMS set lower limits using the $\sqrt{s} = 8$ TeV (20 fb^{-1}) dataset in the

combined dielectron and dimuon channel at $M_S > 4.0$ TeV [34]. This result can be directly measured against the ATLAS $\sqrt{s} = 8$ TeV analysis [1] presented in this thesis.

Table 1.4: Summary of 95% C.L. lower limits on M_S from previous ADD searches.

Collider	Experiment	95% C.L. on M_S
LEP	OPAL	880 GeV
HERA	H1	720 GeV
TEVATRON	D0	1.62 TeV
LHC ($\sqrt{s} = 7$ TeV)	ATLAS	3.22 TeV
LHC ($\sqrt{s} = 8$ TeV)	CMS	4.0 TeV

Contact Interactions

Previous searches for CI have been performed at many different collider experiments. The strongest observed 95% C.L. lower limits set at each collider are described below for the benchmark LL constructive interference model and are summarised in Table 1.5.

At electron–positron collider experiments [35–37] the angular distribution of dilepton and dijet events were investigated. The strongest limit on the CI scale was set at $\Lambda > 10.3$ TeV.

At HERA the differential cross-section ($d\sigma/dQ^2$) was compared to the SM prediction [38] and lower limits were set on the compositeness scale at $\Lambda > 4.0$ TeV.

Dijet angular distributions were studied at the Tevatron to determine lower limits at $\Lambda > 3.1$ TeV [39].

The most recent searches for CI have been performed at the LHC. ATLAS [1, 30] and CMS [34] both used dilepton invariant mass distributions as a search variable. The $\sqrt{s} =$ 8 TeV ATLAS search further included angular information to increase signal-background discrimination. The strongest exclusion limits for CI come from the previous ATLAS non-resonant dilepton analysis using the $\sqrt{s} = 8$ TeV (20 fb^{-1}) dataset [1]. The combined analysis of the dielectron and dimuon channels set lower limits at $\Lambda > 21.6$ TeV. CMS set lower limits at $\Lambda > 18.3$ TeV in the dielectron channel [34].

Table 1.5: Summary of 95% C.L. lower limits on Λ from previous CI searches.

Collider	Experiment	95% C.L. on Λ
LEP	DELPHI	10.3 TeV
HERA	H1	4.0 TeV
TEVATRON	D0	3.1 TeV
LHC ($\sqrt{s} = 8$ TeV)	ATLAS	21.6 TeV
LHC ($\sqrt{s} = 8$ TeV)	CMS	18.3 TeV

"It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."

–Arthur Conan Doyle, Sherlock Holmes

2

Experimental Setup

In order to search for new physics at the highest energy scales achieved in the laboratory, the most sophisticated and innovative accelerator and detector technologies must be employed. This chapter describes the accelerator chain leading up to the LHC followed by the detector design of ATLAS — concentrating on the detection of electrons. A review of data acquisition at ATLAS is provided, highlighting the changes made between the 2012 $\sqrt{s} = 8$ TeV and 2015 $\sqrt{s} = 13$ TeV analysis. Finally a review of the performance is given.

2.1 The Large Hadron Collider

Operated by the European Organisation for Nuclear Research (CERN), the LHC [40] is currently the largest and most powerful particle accelerator ever constructed. With a 27 km circumference, the synchrotron has a design centre-of-mass energy of $\sqrt{s} = 14$ TeV in proton-proton collisions. The LHC is housed in the old Large Electron-Positron (LEP) collider tunnels, located $\sim$ 100 m below the surface of the Swiss Canton of Geneva and crossing the border to the French Pays de Gex.

The LHC requires a series of accelerators to ramp protons to the required input energy of 450 GeV. The process begins by stripping electrons off hydrogen atoms to liberate protons for the first stage of acceleration from Linac 2. Linac 2 is a linear accelerator which provides the Proton Synchrotron Booster (PSB) protons at 50 MeV. The PSB is the first of a series of circular colliders and increases the energy of the protons to 1.4 GeV

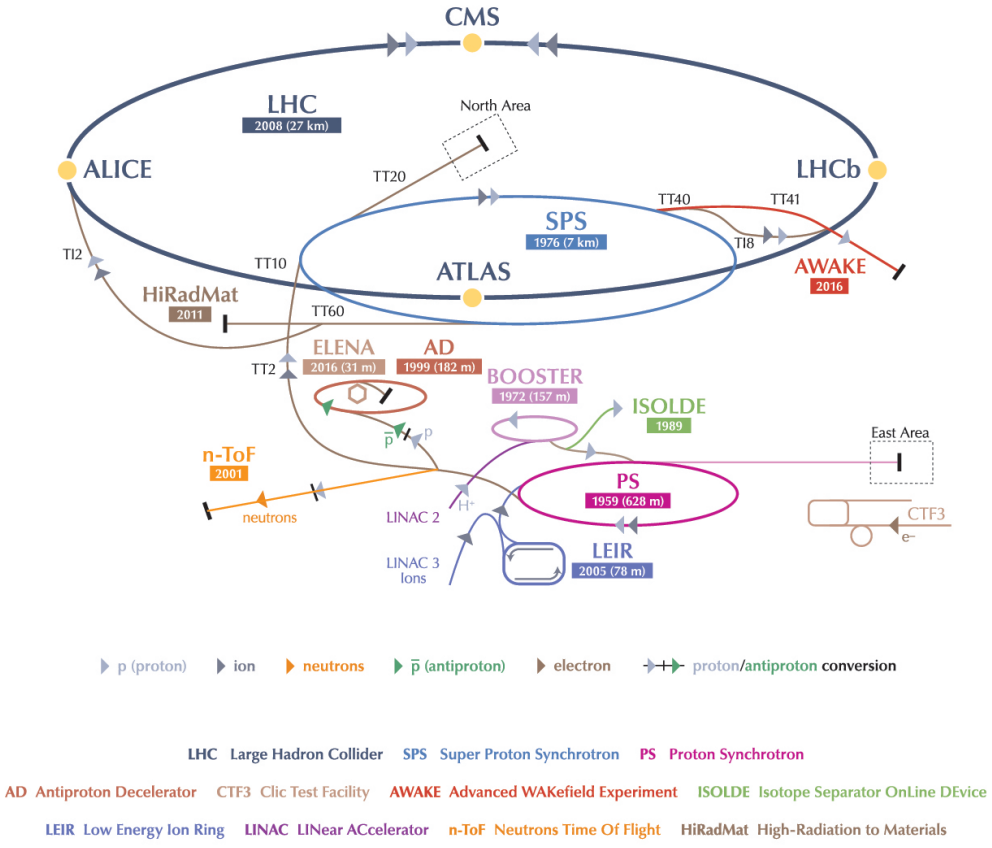


Figure 2.1: LHC accelerating chain and detector positioning [41].

and in turn feeds into the Proton Synchrotron (PS). The PS accelerates protons up to an energy of 25 GeV and injects them into the Super Proton Synchrotron (SPS). The SPS is the penultimate stage of acceleration and provides the LHC with the required 450 GeV protons. Protons circulate the LHC in two oppositely running beams, crossing at four interaction points where the main LHC experiments are housed - including CMS, ALICE and LHCb. Placed at the interaction point closest to the main CERN site is the ATLAS detector. In this thesis, two analyses are presented from 2012 and 2015 where the LHC accelerated the protons to 4 ($\sqrt{s} = 8$ TeV) and 6.5 ($\sqrt{s} = 13$ TeV) TeV, respectively. A schematic of the CERN accelerating complex is given in Figure 2.1.

Many new physics processes are predicted to have low cross-section. As such, a high

instantaneous luminosity is desired to maximise the rate of these possible new interactions. The design instantaneous luminosity of the LHC is 10^{34} cm^{-2}s^{-1} with the beams consisting of bunches of $\sim 10^{11}$ protons. At the end of run 1 (2012), the LHC used 1380 proton bunches at a separation of 50 ns, providing ATLAS with a peak instantaneous luminosity of 7.6×10^{33} cm^{-2}s^{-1}. After long shutdown 1, run 2 commenced in 2015 where the main run used up to 2244 proton bunches at a separation of 25 ns, providing a peak instantaneous luminosity of 5.1×10^{33} cm^{-2}s^{-1}.

2.2 The ATLAS Detector

To maximise sensitivity in a search for an electron-positron (or dielectron) final state, it is paramount to efficiently detect, reconstruct and identify electrons while maintaining as high a purity as possible with the rejection of backgrounds originating from hadrons, electrons from photon conversions, and electrons originating from heavy-flavour decays. This is required over the full detector acceptance and achieved using a combination of complementary detector technologies. The three main subdetector technologies used in ATLAS, when identifying an electron, are the Inner Detector (ID), the Electromagnetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL). The orientation of these subdetectors can be seen in Figure 2.2 and are enveloped by the Muon Spectrometer (MS) at the outermost radii of the detector. The details of the ATLAS sub detectors are extensively discussed in [42], and are briefly described in the following text.

2.2.1 Coordinate System

The coordinate system used to describe the ATLAS detector and the particles emitted from collisions is cartesian or cylindrical depending on convenience. The nominal interaction point (IP) is defined as the origin of each coordinate system, and the beam axis defines the z-axis. The positive x-axis is defined from the interaction point to the centre of the the LHC ring and the positive y-axis is defined as pointing upwards. The azimuthal angle, ϕ, is measured around the beam axis (with origin along the x-axis) and the polar angle, θ, is defined as the angle from the beam axis. The radial distance from the beam pipe is denoted by R. The pseudorapidity is often used to describe a particles trajectory and is

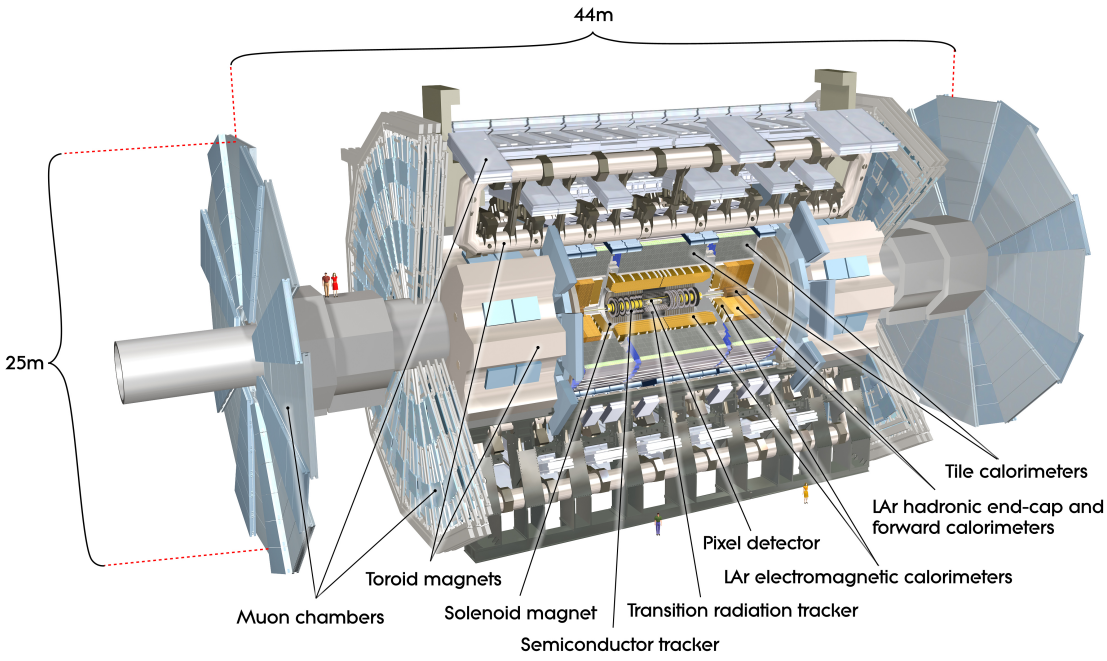


Figure 2.2: Cut away of the ATLAS detector [42].

defined as $\eta = -\ln[\tan(\theta/2)]$, this has the convenience that the differences in η are Lorentz invariant under longitudinal boosts. The distance ΔR in the pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Additionally, transverse quantities such as transverse momentum (p_T) and transverse energy (E_T) are defined in the x-y plane.

2.2.2 Inner Detector

Tracking of charged particles at ATLAS is performed using the ID, which provides robust identification, secondary vertex positioning and spatial track measurements. The ID has coverage within the range $|\eta| < 2.5$ and all ϕ. The design intrinsic resolution of the ID is $\sigma_{p_T}/p_T = 7\%\ p_T \ \oplus\ 1.0\%$ GeV [42]. The ID provides spatial and identification information when linked to energy deposits in the calorimeters or MS tracks. The ID is comprised of separate but complimentary tracking technologies; the Pixel Detector (PD), the Semi-Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT), and in run 2 this also included the Insertable B-Layer (IBL) [43]. Surrounding the ID is a solenoid magnet with a nominal 2 T magnetic field — used to bend the trajectories of charged particles for momentum measurements and charge identification. The orientation

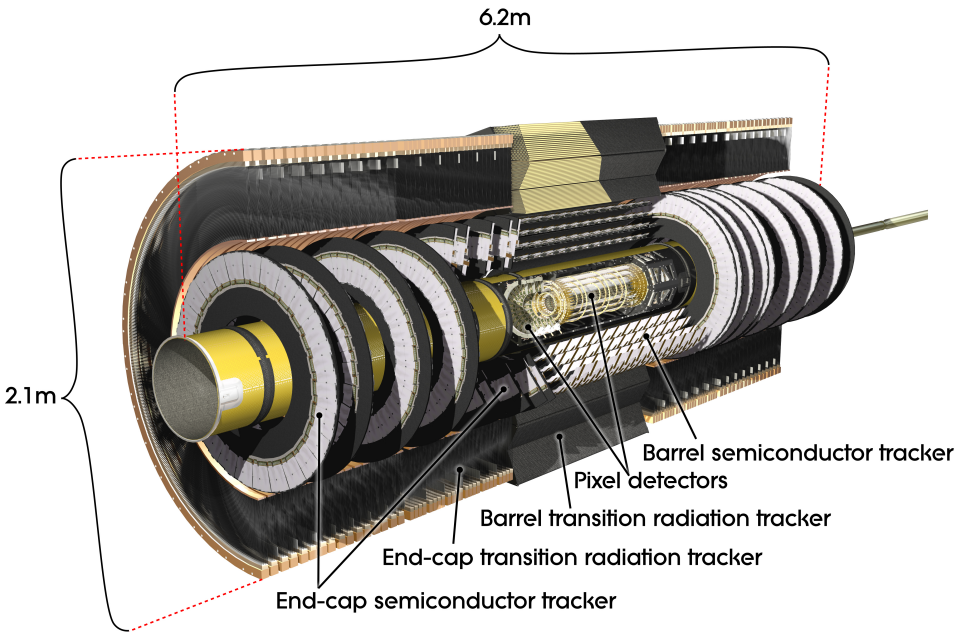


Figure 2.3: Cut away of the ATLAS inner detector [42].

of the ID subdetectors (pre IBL installation) is shown in Figure 2.3. The following sections explain in more detail the technologies used in each layer starting closest to the beam-pipe and working radially outwards.

Insertable B-Layer

During long shutdown 1, the IBL was installed between the PD and the beam-pipe at a radius of 3.3 cm. The IBL is constructed from 14 staves each 64 cm by 2 cm and tilted at an angle of 14° in ϕ. Each pixel is of length 50 µm in ϕ and 250 µm in the z direction. The use of the IBL adds an extra space point to the innermost tacks and aids the identification of secondary vertices.

Pixel Detector

At a radius of 5 cm the PD uses high granularity silicon pixel layers to measure discrete space-points of particle tracks. The Pixel Detector has three components, the barrel and two end-caps each with three layers. Each pixel is of length 50 µm in ϕ and 400 µm in the z direction.

Semi-Conductor Tracker

At a radius of 30 cm the SCT uses the same silicon technology as the PD but in strips of 80 μm by 12.6 cm rather than the high granularity pixels, due to budgetary constraints. The SCT consists of four layers in the barrel region and nine discs in each of the end-caps. Each layer contains two stereo strips arranged at a small-angle (40 mrad) allowing a single space point measurement in r and ϕ for each track.

Transition Radiation Tracker

At a radius of 51 cm the TRT provides the outermost coverage of the ID. The TRT employs a unique detector design which combines tracking measurements with particle identification. The detector is comprised of 4 mm diameter polyamide drift tubes filled with a concentrated xenon gas mixture, chosen for its high efficiency to absorb transition radiated (TR) photons. A gold plated tungsten wire anode is contained within each tube. As highly relativistic particles traverse between the radiator material and the tubes they emit TR, which is absorbed by the active gas mix. The detected signal is then categorised using a low-threshold for track hits (from charged particle ionisation) used in space measurements and a high-threshold for hits (from TR) used in particle identification. Measuring the drift time for the liberated electrons for each hit provides a measurement in $\eta - \phi$ space. The barrel region contains 50000 straws parallel to the beam axis with the anode electrically divided into two halves at $\eta = 0$. Barrel readouts are positioned at both ends of the TRT. In the end-cap, the tubes are arranged radially in wheels, containing 160000 tubes each and a single readout per tube at the outer radius. Typically 36 hits per track are measured in the TRT.

The performance of the TRT when discriminating signal electrons over charged pions can be found in Section 3.2.

2.2.3 Calorimetry

Energy measurements are performed in the ATLAS calorimeters over the range $|\eta| <= 4.9$. There are two calorimeter subdetectors optimised to measure the energy of different physics processes. The ECAL is designed to measure the energy of electrons and photons, while the HCAL detects energy deposits from hadrons. To accurately measure the energy of a par-

ticle, the respective calorimeters must contain all electromagnetic and hadronic showers. This is also necessary to stop showers of particles entering the MS and being miss reconstructed as muons (punch-through), and as such was taken into the design considerations for the calorimeter depth.

Electromagnetic Calorimeter

The ECAL is a lead/liquid argon (LAr) sampling calorimeter designed to measure electromagnetic particle energy deposits and provide particle identification. Electrons interact with the lead absorbing material through the Bremsstrahlung process resulting in an electron and photon. Photons interact with the lead in a pair production mechanism resulting in an electron and positron. These processes create a shower of particles which ionise the active liquid argon sampling material, placed between alternating layers of absorber. The energy deposited in the calorimeter is then measured from the voltage induced by the ionisation. Figure 2.4 shows the accordion structure of the ATLAS ECAL.

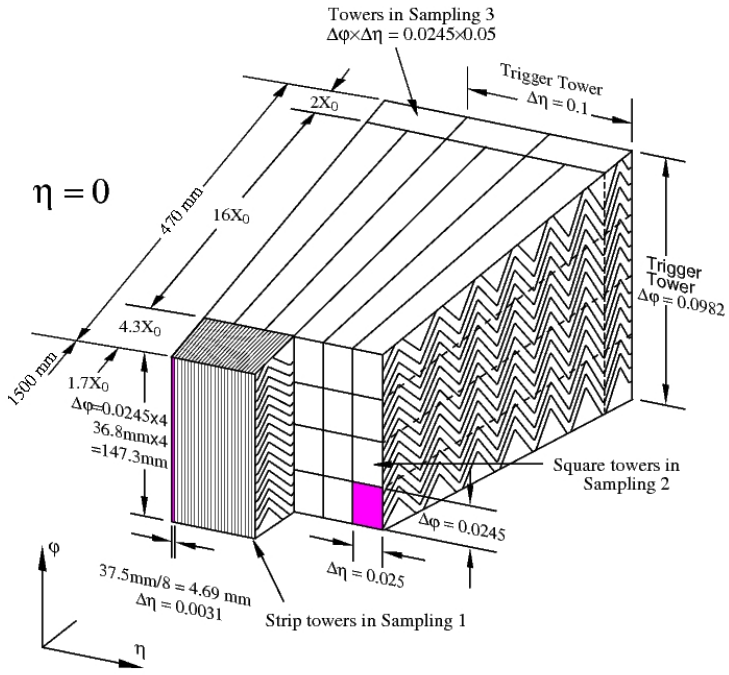


Figure 2.4: Accordion structure of the ATLAS Electromagnetic Calorimeter [42].

The central ECAL has coverage of $|\eta| < 3.2$ and is split into the barrel and two end cap regions. The design intrinsic resolution of the ECAL is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ GeV [42].

Using lead absorber layers interleaved between the active layers in an accordion structure ensures a complete ϕ symmetry without azimuthal cracks. There are three sampling layers for $|\eta| < 2.5$ and two layers for $2.5 < |\eta| < 3.2$. The first sampling layer is constructed in narrow η strips allowing detection of π^0 decays, where the two photon final state could be reconstructed as a single electromagnetic particle at lower granularity. The calorimeter is preceded by a pre-sampler as a first layer of active material at $|\eta| < 1.8$ to recover energy lost in support material. The total depth of the ECAL is > 22 radiation lengths[1] (X_0) in the barrel and > 24 X_0 in the end-caps. The Forward LAr calorimeter extends the calorimeter coverage to the range $3.1< |\eta| <4.9$ and uses LAr as the active material with copper and tungsten as the absorber material.

The performance of the ECAL when discriminating signal electrons over jets can be found in Section 3.2.

Hadronic Calorimeter

The HCAL is comprised of two technologies, the Tile Calorimeter (Tile) in the barrel with coverage at $|\eta| < 0.8$ and end-caps at $0.8 < |\eta| < 1.7$, and the LAr Hadronic Endcap (HEC) extending coverage to $1.5 < |\eta| < 3.2$. The Tile uses alternating layers of scintillator as the active material and steel as the passive absorber. The HEC again uses LAr as the active medium and copper as the passive material. The HCAL is designed to fully stop all hadronic showers and has a depth of ~ 9.7 interaction lengths[2] (λ) in the barrel and ~ 11 λ in the end-caps. The HCAL is used in electron selection by checking if there is any shower leakage from the ECAL into the HCAL which could be indicative of a mis-identified jet.

2.2.4 Muon Spectrometer

The MS is located at the outermost radii of ATLAS. It is designed to measure the momentum of charged particles exiting all other layers of the detector. The ATLAS toroid magnet is used to bend charged trajectories for momentum measurements. The MS has four sub systems: Monitor Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers (RPC), and Thin Gap Chambers (TGC). The MDT measure momentum

[1] A radiation length corresponds to the distance over which, on average, an electron loses (1 - 1/e) of its initial energy.

[2] The average distance a hadronic particle travels inside the material before an inelastic interaction.

at $|\eta| < 2.0$ and the CSC at $2.0 < |\eta| < 2.7$. The triggering of muons in the barrel region $|\eta| < 1.5$ is performed with the RPC and extended by the TGC up to $|\eta| < 2.4$.

2.2.5 Magnet Systems

ATLAS uses two magnetic systems to bend charged particles, the solenoid and the toroid. The solenoid is housed between the ID and the ECAL, sharing the ECAL cryostat. This reduces the dead material of the solenoid system to 0.66 X_0. Operating at 4.6 K, a 2 T magnetic field is produced for momentum measurements in the ID. The toroid is split into three sections: the barrel and two end caps. In the barrel, eight coils equally spaced in ϕ generate a 3.9 T magnetic field. The end-cap toroid system provides an additional field in the forward region of 4.1 T. The toroid provides the bending power used in muon momentum and charge measurements.

2.2.6 Trigger System

The LHC is designed to deliver beams of proton bunches at a crossing rate of 40 MHz; too large to be reconstructed or recorded. The ATLAS trigger system was developed to identify a subset of potentially interesting events from detector signatures and ensure that these were recorded. In run 1, the trigger system consisted of three discrete levels: Level-1 (L1), Level-2 (L2), and the Event Filter (EF). The L1 is a hardware based trigger where L2 and EF collectively form the software based High-Level Trigger (HLT). In run 2 the L2 and EF merged to become a single software trigger - the HLT.

Level-1 Trigger

The first layer of the trigger system is the fully hardware based L1 trigger. Designed to reduce event rates down to 75 kHz in run 1 and upgraded to 100 kHz for run 2. Using reduced granularity input from calorimeter deposits and MS hits, "regions of interest" (RoI) are selected in $\eta - \phi$ cones if a threshold energy is reached. The target latency of the system is 2 μs, most of which facilitates the transmission of data of an accepted event to a nearby computer farm for further analysis. These RoI seed the HLT; if the L1 trigger fails to find a RoI - the event is rejected.

Run 1 - Level-2 and Event Filter

In run 1, the second layer of the trigger system was the software based L2 trigger. Using the RoI found at L1, the L2 uses ID tracking information in addition to the full-granularity RoI data to reduce the event rate to below 5 kHz. The L2 trigger employs reconstruction algorithms of shower shapes, track-cluster matching and E_T thresholds to make a decision if the event is kept. A decision is made and read out in 40 ms, if the requirements contained in trigger chain are met the event is passed on to the EF, if not it is rejected.

The final stage of triggering was performed at the EF. Again using the full detector granularity, if an event meets another set of increasingly strict requirements it is stored for offline analysis. With a 1 s latency the EF can fully reconstruct events and save them to disk at a rate of 400 Hz.

Run 2 - High Level Trigger

In run 2 the L2 and EF levels were merged together into a collective HLT trigger. The motivation for this change stems from the separate L2 and EF computing farms used in run 1, which as a consequence caused time wasted transferring data, and even some processing duplication. With a 100 kHz input from L1 the HLT is designed to reduce this rate to 1 kHZ, again using the full detector granularity and software based algorithms to determine which events to save for offline analysis.

2.2.7 Performance

In the $\sqrt{s} = 8$ TeV 2012 data taking period, a total of 21.3 fb^{-1} of data was recorded by ATLAS over roughly eight months. In the $\sqrt{s} = 13$ TeV 2015 data taking period, a total of 3.9 fb^{-1} of 25 ns data was recorded by ATLAS over approximately five months. The rate at which this was recorded can be seen in Figures 2.5 and 2.6 for the $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV runs, respectively. Only good quality data is desired for offline analysis and a series of recommended good runs lists (GRLs) are available to verify certain detector components were functioning nominally during each run. As such the quoted dataset used for analysis is a subset of the ATLAS recorded dataset.

In each bunch crossing multiple protons can collide, this is known as in-time pile-up. This makes the identification of the event of interest extremely challenging as an order of

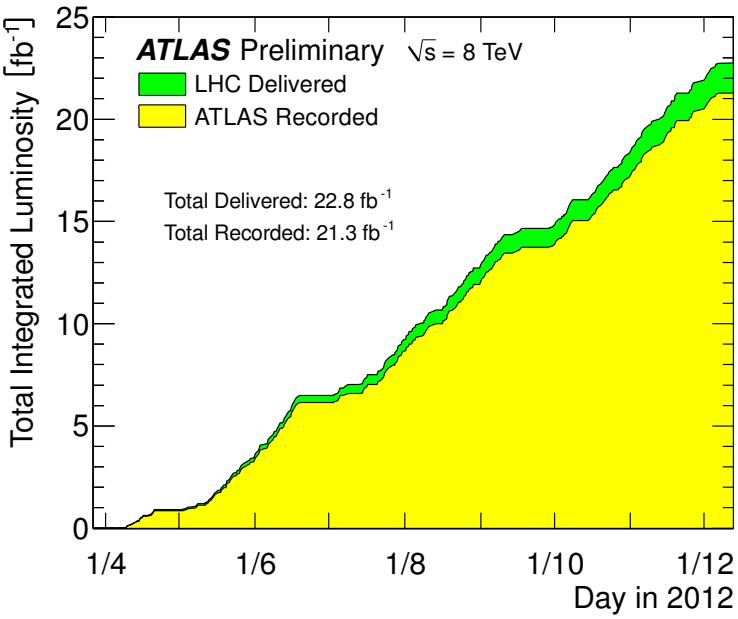


Figure 2.5: ATLAS integrated luminosity accumulation in 2012 [44].

10^3 tracks are produced. The parameter μ is defined as the number of particle interactions per bunch crossing. The average number of interactions per bunch crossing is denoted as $<\mu>$. The μ profile for data is shown in Figure 2.7 and 2.8 for the $\sqrt{s} = 8$ TeV and $\sqrt{s} = 8$ TeV runs, respectively. The lower run 2 $<\mu>$ is due to the lower instantaneous luminosity. With run 2 continuing in 2016 this is expected to increase as the delivered beams are focused to smaller areas as they cross.

2.3 Collected Datasets

The data used in the ADD search consists of the 2012 dataset collected by the ATLAS detector during LHC pp collisions at $\sqrt{s} = 8$ TeV, corresponding to 20.3 fb^{-1}. The data used in the CI search consists of the 2015 25 ns dataset collected during LHC pp collisions at $\sqrt{s} = 13$ TeV, corresponding to 3.2 fb^{-1}. After data has been recorded from the detector, events can be reconstructed, pattern recognition performed to identify particle types and momenta and energies calculated.

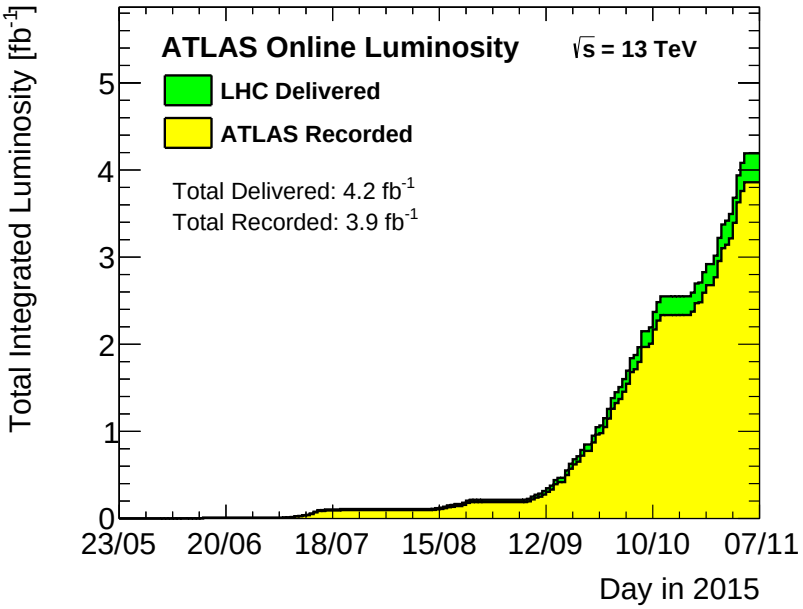


Figure 2.6: ATLAS integrated luminosity accumulation in 2015 [45].

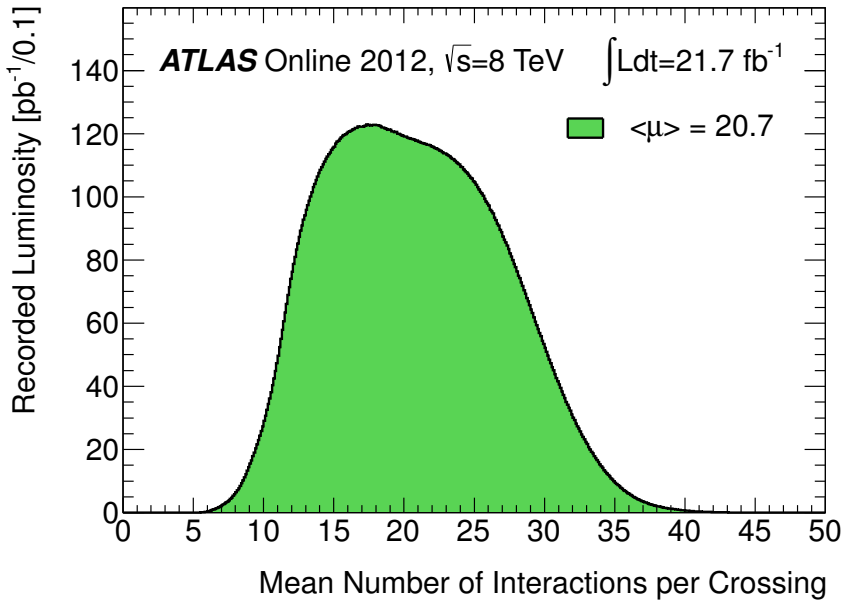


Figure 2.7: ATLAS pile-up in 2012 dataset [44].

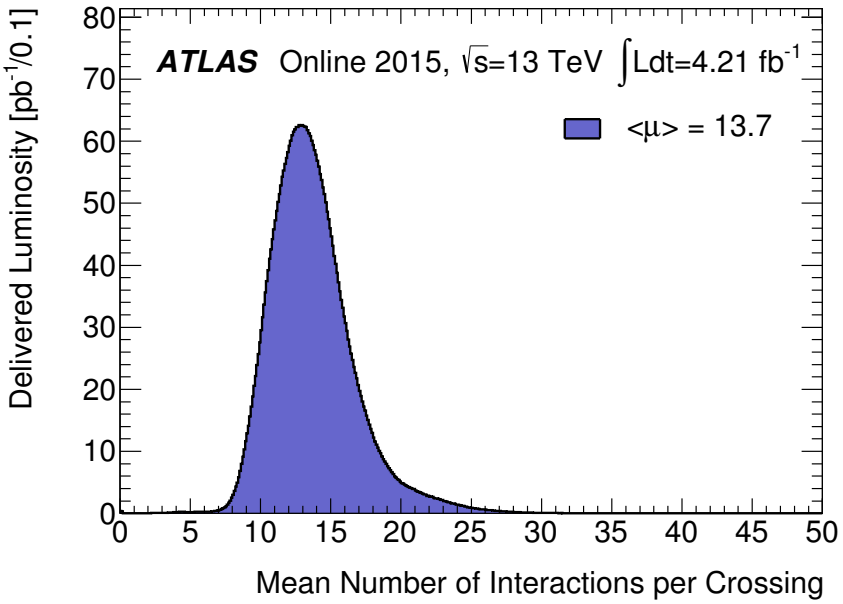


Figure 2.8: ATLAS pile-up in 2015 dataset [45].

"Never send a human to do a machine's job."

–Agent Smith

3

Electron Reconstruction

After electronic signals are recorded in ATLAS they are reconstructed to determine the original particle type. Particles cannot be identified absolutely, only inferred from matches to criteria that build up an object definition. This chapter reviews how electrons are reconstructed at ATLAS using ID tracking, ECAL shower shape, HCAL activity, and isolation measurements. Firstly, a general overview of how electrons are expected to interact with the ATLAS detector is given with a discussion of energy clustering and track-cluster matching algorithms. The criteria used to separate signal electrons from numerous backgrounds is discussed. Details of the various purities of electron identification provided by the central ATLAS electron/photon group are reviewed. Finally the use of additional isolation criteria is described.

3.1 Reconstruction Algorithms

As electrons are electrically charged they interact with the inner detector leaving a trail of space points which can be traced into the ECAL where an electron is expected to deposit all of its energy. The first step in reconstructing an electron is the clustering of ECAL energy deposits to a single source. Secondly, reconstructed tracks are created from space point measurements in the ID that can be matched to the clustered energy deposits. The method used to calibrate the signal measured in the detector to the energy of the incident particle are discussed in Section 4.1.2.

Energy Clustering

Energy deposits in the ECAL are summed over each of the three layers plus pre-sampler to form towers. This is performed using the same granularity as that of the second ECAL layer. If any overlap occurs, the energy is attributed equally to each of the participating towers. A "sliding window" algorithm is applied to these towers, summing cells within a fixed-size rectangular window (3×5) and adjusting the position of the window such that the transverse energy is a local maximum [46]. A seed cluster is made if the maxima has $E_T > 2.5$ GeV. Using MC simulation and a pre-selection cut on the angular separation between the generated and reconstructed electrons, real sources are found to be reconstructed with around 95% efficiency for electrons with $E_T > 7$ GeV, 99% for $E_T > 15$ GeV and 99.9% for $E_T > 45$ GeV [47]. For each seed, a RoI of cone-size $\Delta R = 0.3$ around the cluster barycentre is defined and used in the track reconstruction algorithm.

Track Reconstruction

Track reconstruction uses a Kalman filter algorithm to connect space-point hits into a track [48]. To begin the algorithm, three hits from the silicon detectors, which are compatible with a charged particle trajectory to a defined ECAL cluster cone are required. These hits are propagated radially outwards into the TRT. This process is then repeated, starting with the TRT, which is important for determining tracks originating from photon conversions (pair production of two electrons). All tracks with $p_T > 0.4$ GeV and $|\eta| < 2$ are tested. If the difference in ΔR between the extrapolated track and the seed cluster is > 0.01, the matching is rejected. The electron cluster is then rebuilt using towers of cells of 3 $\times$7 dimension in the central ECAL second layer to define an electron object.

3.2 Electron Identification

Electron identification criteria are used to reject the maximum amount of backgrounds while keeping signal efficiency high. These backgrounds occur from multiple sources and specific criteria are used to minimise these contributions. The full set of variables used by the ATLAS electron/photon group are presented in Table 3.1 for shower shape and Table 3.2 for tracking variables. A few particularly pertinent variables are discussed below.

Table 3.1: The shower shape variables associated with electron identification [49].

Type	Description	Name
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| < 0.8$ or $\|\eta\| > 1.37$)	R_{had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\| > 0.8$ and $\|\eta\| < 1.37$)	R_{had}
Back layer of EM calorimeter	Ratio of the energy in the back layer to the total energy in the EM accordion calorimeter	f_3
Middle layer of EM calorimeter	Lateral shower width, $\sqrt{(\sum E_i\eta_i^2)/(\sum E_i) - ((\sum E_i\eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	$W_{\eta,2}$
	Ratio of the energy in 3×3 cells over the energy in 3×7 cells centred at the electron cluster position	R_ϕ
	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centred at the electron cluster position	R_η
Strip layer of EM calorimeter	Shower width, $\sqrt{(\sum E_i(i - i_{max})^2)(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	ω_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
	Ratio of the energy in the strip layer to the total energy in the EM accordion calorimeter	f_1

Table 3.2: The tracking variables associated with electron identification [49].

Type	Description	Name
Track quality	Number of hits in the B-layer (discriminates against photon conversions)	n_{Blayer}
	Number of hits in the pixel detector	n_{pixel}
	Number of total hits in the pixel and SCT detectors	n_{Si}
	Transverse impact parameter	d_0
	Significance of transverse impact parameter defined as the ratio of d_0 and its uncertainty	σ_{d_0}
	Momentum lost by the track between the perigee and the last measurement point divided by the original momentum	$\Delta p/p$
TRT	Total number of hits in the TRT	N_{TRT}
	Ratio of the number of high-threshold hits to the total number of hits in the TRT	F_{HT}
	Likelihood probability based on transition radiation in the TRT	TRTPID
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track ($\lvert\Delta\eta_1\rvert < 0.01$)	$\Delta\eta_1$
	$\Delta\phi$ between the cluster position in the middle layer and the extrapolated track ($\lvert\Delta\phi_2\rvert < 0.02$)	$\Delta\phi_2$
	$\Delta\phi$ between the cluster position in the middle layer and the extrapolated track, where the track momentum is rescaled to the cluster energy before extrapolating the track to the middle layer of the calorimeter	$\Delta\phi_{Res}$
	Ratio of the cluster energy to the track momentum	E/p
Conversions	Veto electron candidates matched to reconstructed photon conversions	isConv

Photons do not leave tracks in the Inner Detector so an electron candidate requires a track leading to the energy deposit. A photon conversion however could produce an electron-positron pair, as such a minimum requirement on the number of innermost silicon hits is required to minimise this background (n_{Blayer}).

The objects built by the electron reconstruction algorithms are not always electrons from the hard scatter process or even real electrons. Fake electron candidates can come from hadronic jets as well as real background electrons from semi-leptonic heavy flavour decays. For every isolated electron with $E_T \sim 1$ TeV there are typically 10^3 jets in proton-proton collisions. Typically jets consist of ~ 15 particles, however a jet consisting of collimated π^0s and a $\pi^\pm$ could provide a track and electromagnetic shower ($\pi^0 \to \gamma\gamma$) which can be wrongly reconstructed as an electron.

As discussed in Section 3.2 the TRTs novel design is motivated by electron-pion discrimination through the emission of transmission radiation (TR). The difference in Lorentz factors (γ factor) of electrons and pions causes different numbers of high energy TR emissions and can be used as a discriminant (F_{HT}). The ability to separate these particles is shown in Figure 3.1 for the expected and measured electron and charged pion high energy TR probability as a function of Lorentz factor.

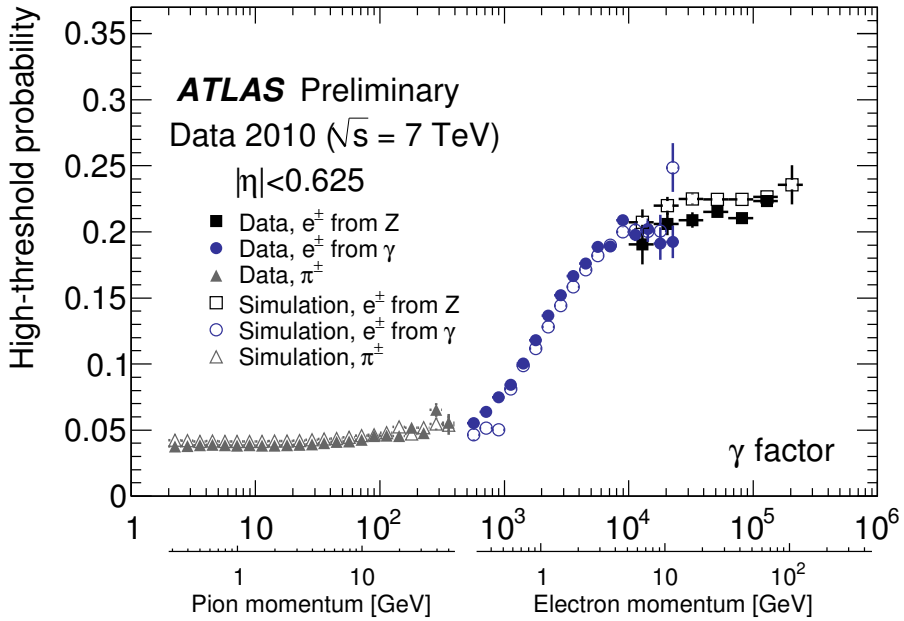


Figure 3.1: Lorentz factor used in electron-pion discrimination in the TRT [50].

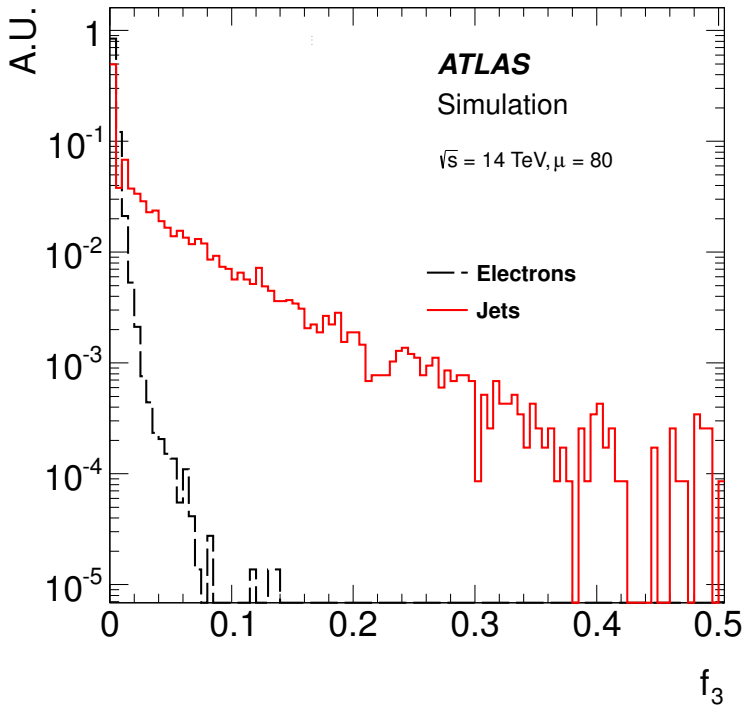


Figure 3.2: Expected shower shape variable f_3 used in electron-jet discrimination, normalised to arbitrary units (A.U.) [51].

An electron is expected to deposit most of its energy in the first two layers of the ECAL. As jets interact with the ECAL and the HCAL, they can be rejected using the segmented calorimeter measurements. An example of using segmentation to separate electrons from jets is shown in Figure 3.2 for the simulated energy fraction measured in the ECAL third layer compared to the total ECAL (f_3).

An electron is expected to deposit energy within a smaller area than a jet. As such the lateral shower shape measurements can also be used to discriminate from jets. Using the granularity of the second layer of the ECAL, ratios of different sized cell windows can be constructed from energy deposits. As the electron energy is mostly contained in the subset of cells the ratio is expected to be unity, while jets take a wider range of values due to the wider energy deposits. An example of this can be seen in Figure 3.3 using the ratio of 3×7 and 7×7 energies (R_η). Another lateral shower shaped variable employed is the lateral width ($W_\eta, 2$) calculated with a window of 3×5 cells using the energy weighted sum over all cells. The effectiveness of this variable is shown in Figure 3.4.

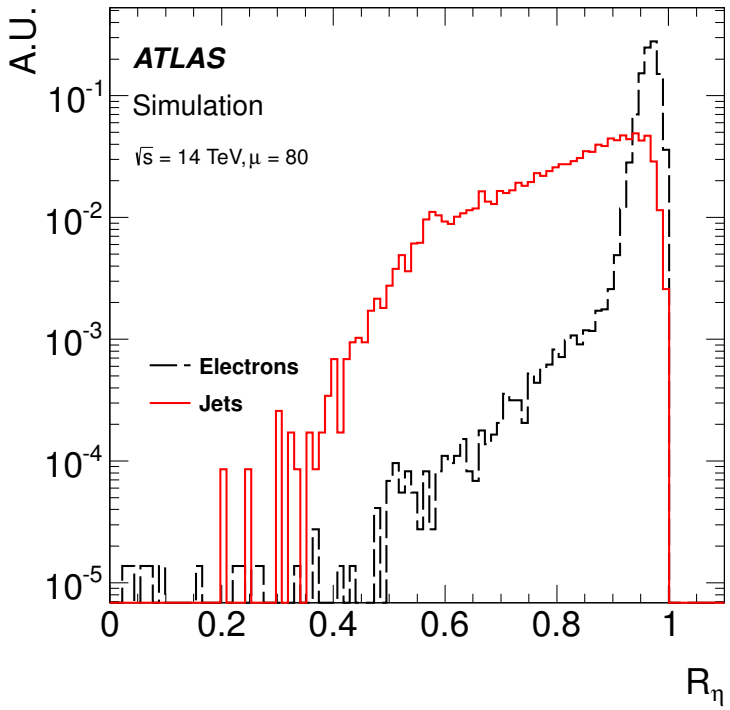


Figure 3.3: Expected shower shape variable R_η used in electron-jet discrimination, normalised to arbitrary units (A.U.) [51].

To reduce the contamination of signal electrons from the hard scatter process with electrons from heavy flavour decays, the transverse impact parameter, d_0, is used. This is the point of closest approach with respect to the beam line. Due to the delayed decay of heavy flavour mesons, this parameter is expected to be larger than that measured for signal electrons. A study was performed on the signal efficiency and background rejection of an additional cut on this parameter in the $\sqrt{s} = 13$ TeV analysis as described in Section 5.2.1.

The stringency of the cuts on the parameters used in electron identification give varying signal efficiency and background rejection. The three working points *loose*, *medium* and *tight* use increasingly strict variable cuts, providing higher purities of signal electrons and are chosen on an analysis by analysis basis depending on the specific requirements. The definition of *medium* was replaced by *medium++* in run 1 after a re-optimisation in the selection criteria making the identification tighter including a selection based on the number of inner most silicon hits.

The variables used in each of the identification working points is listed in Table 3.3. The $\sqrt{s} = 8$ TeV analysis uses rectangular cuts on these variables [47], however the $\sqrt{s} =$

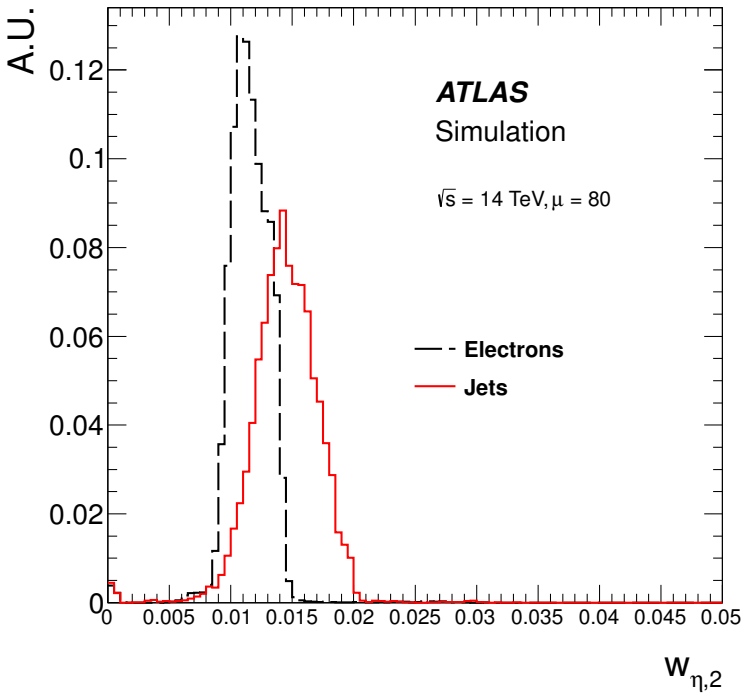


Figure 3.4: Expected shower shape variable $W_{\eta,2}$ used in electron-jet discrimination, normalised to arbitrary units (A.U.) [51].

13 TeV analysis used a likelihood approach [49]. A study into the effect on signal selection efficiency and background rejection is described in Section 5.2.2 for the $\sqrt{s} = 13$ TeV analysis.

3.3 Electron Isolation

A further method in which to reject jets mis-identified as electrons and real electrons is the use of additional isolation requirements. This can be based on the energy deposits in the calorimeter or the the track density in the inner detector.

In run 1, the calorimetric isolation variable $E_T cone X$ is defined as the sum of the transverse energy deposited in the calorimeter cells in a cone ΔR of size X/100 around the electron barycentre, excluding the electron cluster contribution. This variable is used to measure the amount of additional activity there is around the electron. The $\sqrt{s} = 8$ TeV dielectron event selection used an E_T dependent requirement on $E_T cone20$, constructed to maintain a 99% signal efficiency [1].

Table 3.3: Variables used in *loose*, *medium* and *tight* identification definitions [47, 49].

	8 TeV Cut Based			13 TeV Likelihood Based		
Name	*loose*	*medium++*	*tight*	*loose*	*medium*	*tight*
R_{Had1}	✓	✓	✓	✓	✓	✓
R_{Had}				✓	✓	✓
f_3		✓	✓	✓	✓	✓
$W_{\eta 2}$	✓	✓	✓	✓	✓	✓
R_η	✓	✓	✓	✓	✓	✓
R_ϕ				✓	✓	✓
ω_{stot}	✓	✓	✓			
E_{Ratio}	✓	✓	✓	✓	✓	✓
f_3				✓	✓	✓
n_{Blayer}		✓	✓		✓	✓
n_{Pixel}	✓	✓	✓	✓	✓	✓
n_{Si}	✓	✓	✓	✓	✓	✓
d_0		✓	✓	✓	✓	✓
σ_{d_0}				✓	✓	✓
$\Delta p/p$				✓	✓	✓
n_{TRT}		✓	✓			
F_{HT}		✓	✓			
TRTPID				✓	✓	✓
$\Delta\eta_1$	✓	✓	✓	✓	✓	✓
$\Delta\phi_2$			✓			
$\Delta\phi_{Res}$				✓	✓	✓
E/p			✓			
isConv			✓			

In run 2, this isolation variable evolved into topo$E_T coneX$. This only sums the energy deposits within the defined cone if it is selected by the topological clustering algorithm which groups neighbouring cells in the ECAL as long as the signal in each cell is significant with respect to noise [46]. This suppresses noise and protects the measurement from out-of-time pile-up caused by collisions in different bunch crossings. If electronic signal in the calorimeter remains from energy deposits from previous events (out-of-time pile-up) this can be incorrectly associated to the hard scatter process of interest.

The track isolation variable, $p_T coneX$, is defined as the sum of the transverse momentum of the tracks within a cone ΔR of size X/100 around the electron track. The tracks considered in the sum must come from the primary vertex associated to the electron track and be of good quality, i.e. they must have at least nine silicon hits, one of which should

be in the innermost silicon layer. $p_T varconeX$ has a variable cone size depending on the the leptons p_T [52].

The *loose* isolation working point is chosen for the final $\sqrt{s} = 13$ TeV dielectron event selection, and is designed to maintain a 99% signal efficiency for the combined track and calorimeter isolation. The track-based isolation puts a requirement on $p_T varcone20/p_T$, while the calorimeter-based isolation puts a requirement on $topoE_T cone20/E_T$. The *loose* working point provided a rejection of up to 30% of b/c flavoured meson decays in $t\bar{t}$ events throughout the available invariant mass range [4].

"The generation of random numbers is too important to be left to chance."

–Robert R. Coveyou

4

Background and Signal Modelling

In order to test data against predictions, the contribution from relevant processes have to be estimated accurately. This chapter details the methods of background and signal estimation. Firstly a description of the event simulation and how accurate the Monte-Carlo estimates are is given, followed by the details of the programs used in these searches. The corrections applied to the simulated backgrounds are discussed where there is known disagreement with data. Finally a review of the data-driven technique employed in the W+jets and multi-jet background is presented.

While this thesis contains two different collision energies and two BSM signatures, the background processes have a large overlap. The modelling of SM processes used in both analyses are discussed simultaneously where possible.

4.1 Event Simulation

The accurate simulation of SM backgrounds and BSM processes is fundamental when searching data for new physics. This can be achieved through simulation using Monte Carlo (MC) processes. This section will discuss the simulation of events and what corrections are applied to MC in order to accurately describe data.

In ATLAS, the simulation of events consists of four steps: event generation, detector response using the GEANT4 [53] software framework, digitisation and reconstruction. The output of digitisation has the same format as the data retrieved from the ATLAS detector,

thus the reconstruction step is common for both recorded and simulated data.

4.1.1 Event Generation

An event generator simulates a physics process from a single interaction vertex and provides the 4-momenta and particle type of the final state particles. It is important to predict these events as accurately as possible, however, due to the perturbative nature of these calculations this poses a theoretical and computational challenge. The simplest cross-section calculation is from the leading-order (LO) diagram, which has the minimum number of vertices in a reaction (two). The LO calculation is the largest contribution to the total cross-section. Additional diagrams containing extra vertices will increase the accuracy of the calculation and better describe the data. These higher-order corrections come from additional electroweak or QCD processes. The higher-order electroweak corrections have contributions from real initial/final state photon radiation and electroweak loops. The higher-order QCD corrections have contributions from real gluon radiation or virtual quark/gluon loops. These high-order processes are named next-to-leading-order (NLO) for diagrams with up to three vertices, next-to-next-to-leading-order (NNLO) when including diagrams with up to four vertices, and so on.

The accuracy of different order generators compared to data when selecting high mass DY events is displayed in Figure 4.1. The lower ratio plot shows MC predictions are consistent with the shape of the measured dielectron invariant mass distribution.

4.1.2 Corrections

MC descriptions of events are known to be imperfect, either due to the mis-modelling of data, run conditions, detector defects or reconstruction inefficiencies. As such data derived corrections are applied to MC samples in both analyses.

Pile-Up

During data-taking, the number of proton-proton interactions is measured for each bunch crossing by counting the number of vertices. MC samples are generated across a broad range of pile-up values which are weighted according to the average number of interactions per bunch crossing as measured in data.

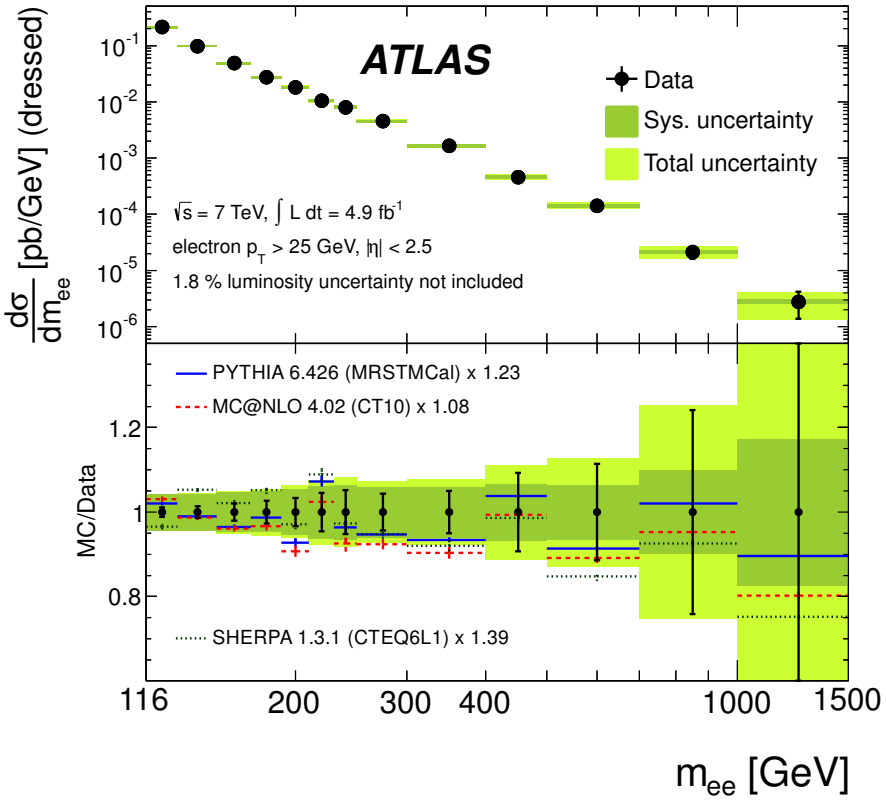


Figure 4.1: Measured Drell–Yan differential cross-section within the fiducial region (electron $p_T >$ 25 GeV and $|\eta| < 2.5$). In the lower plot, the measurement is compared to the predictions of the PYTHIA, MC@NLO and SHERPA MC generators including their statistical uncertainties [54].

Energy Scale and Resolution

The signal from an ECAL cluster is calibrated to correctly account for the energy of the electron that initiated the shower. This is known as the energy scale and is determined through test beam measurements. Due to energy losses in upstream material, lateral leakage (energy not contained in the cluster cells) and energetic showers which are not fully contained within the ECAL depth, in-situ η and E_T dependent corrections are derived from MC and data. The scaling is derived from dielectron pairs from Z decays using the known mass peak. After these corrections at $\sqrt{s} = 8$ TeV the energy scale is known to an accuracy of 0.03 – 0.2% (0.3 – 2.3%) for electrons with $E_T = 40$ GeV (200 GeV), depending on pseudorapidity [55].

$Z \to ee$ events are used to correct MC to the measured resolution in data. The energy of simulated electrons is smeared such that it matches the Z peak. A resolution smearing correction of 0.8% was determined in the barrel [55].

Since the 2015 data used the scale and smearing corrections from extrapolating the $\sqrt{s} = 8$ TeV data and not direct $\sqrt{s} = 13$ TeV measurement an additional uncertainty (typically 0.2 – 0.3% on the electron scale at 40 GeV E_T) is assigned to these corrections.

Reconstruction Efficiencies

The total efficiency of selecting an electron is the product of electron reconstruction, identification, isolation and trigger efficiencies. In ATLAS, efficiencies for electrons with $E_T > 25$ GeV are measured in $Z \to ee$ events using the "tag-and-probe" method to obtain a clean sample of electrons [47]. If an event contains an electron which passes a series of tight selection criteria it is used to tag the event. The second electron, which passes a series of looser selections and kinematic criteria (to increase purity of true Z decays) can be used to probe the efficiency of the selection of interest.

Scale factors (SF) for selection efficiencies are calculated as the efficiency observed in data divided by that in MC. This can be measured over a broad range of η and E_T values to create a map of SF to be applied to MC. All SF are found to be within 1 and 2% of unity across η and E_T in the $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV analysis, respectively [47,56].

4.2 Standard Model Backgrounds

The main background to either ADD graviton events or contact interactions (CI) decaying to two electrons is the irreducible Drell–Yan (DY) process. Another smaller irreducible contribution comes from the Photon-Induced (PI) processes. Other non-negligible but reducible backgrounds arise from top quark and diboson (WW, WZ and ZZ) decays with at least two electrons in the final state. These backgrounds are simulated using MC. Backgrounds originating from miss identified jets, typically multi-jet and W+jets events, are estimated from a data-driven method as described in Section 4.2.5.

4.2.1 Drell–Yan Background

The largest background to these analyses is the irreducible DY process. The leading order process is shown in Figure 1.3. The DY estimate uses a large un-binned sample and a series of mass-binned samples. This ensures enough events in both the Z-peak region, which is used to normalise the background prediction to data, and at high masses where signal could be present. The un-binned sample is generated at a dielectron truth mass of > 60 GeV, an upper limit of 120 GeV is imposed such that it is smoothly stitched to the mass binned samples beginning at 120 GeV. In the $\sqrt{s} = 8$ TeV analysis the mass-binned samples are split into 15 MC truth dilepton mass regions between 120 GeV and > 3000 GeV. In the $\sqrt{s} = 13$ TeV analysis, the mass-binned samples are split in to 19 MC truth dilepton mass regions between 120 GeV and > 5000 GeV.

Both analyses use the NLO event generator programme POWHEG [57] (implementing the CT10 [58] PDF) and the programme PYTHIA8 [59] for event showering. A correction (k-factor) is applied to weight the cross-section from NLO to NNLO in QCD and to NLO in EW. These corrections are dilepton mass-dependent. The $\sqrt{s} = 8$ TeV analysis derived this NNLO k-factor using the programme FEWZ [60] using the MSTW2008 NNLO PDF [61]. The $\sqrt{s} = 13$ TeV analysis used the programme FEWZ and the CT14NNLO PDF [62].

4.2.2 Photon-Induced Background

Another source of irreducible background arises through photons originating from incoming quarks decaying to electrons. These PI events are not accounted for in the generation of DY events and so are described with a dedicated sample in the $\sqrt{s} = 8$ TeV analysis and a mass-dependent correction in the $\sqrt{s} = 13$ TeV analysis. The leading order PI contribution can be seen in Figure 4.2.

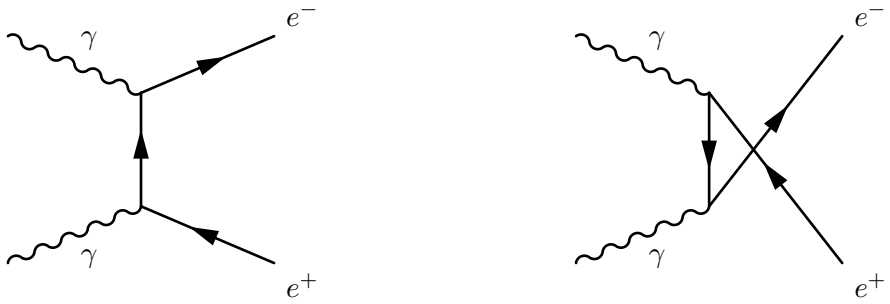


Figure 4.2: Leading-order Feynman diagrams for Photon-Induced process.

The $\sqrt{s} = 8$ TeV analysis estimated this contribution with the PYTHIA8 programme [59] at LO using the MRST2004QED PDF [63]. To ensure adequate statistics, PI samples are generated in 4 truth invariant mass bins of 60 – 200 GeV, 200 – 600 GeV, 600 – 1500 GeVand 1500 – 2500 GeV.

4.2.3 Top Quark Background

The main reducible background arises from decays involving top quarks. This background consists of the $t\bar{t}$ and Wt (a single top in association with a W boson) processes. As top quarks decay into a W boson and a bottom quark, if the W further decays into an electron and a neutrino, a resulting dielectron final state can be observed, as shown in Figure 4.3 and Figure 4.4 for $t\bar{t}$ and Wt events, respectively.

The $\sqrt{s} = 8$ TeV analysis uses the programme MC@NLO 3.41 [64] with NLO CT10 PDF [58] to generate matrix elements. The programme JIMMY 4.31 [65] is used to simulate parton interactions and HERWIG [66] for the underlying event and parton showers.

The $\sqrt{s} = 13$ TeV analysis uses the programmes POWHEG [57] with NLO CT10 PDF to generate matrix elements, and PYTHIA6 [67] for the event showering.

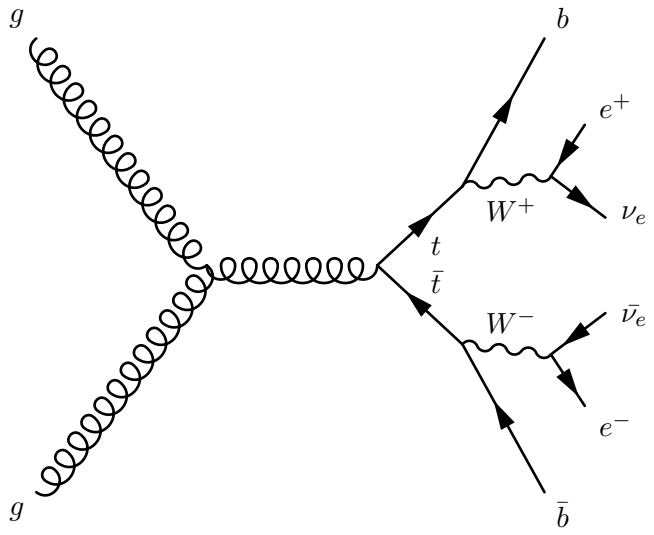


Figure 4.3: Leading-order Feynman diagram for $t\bar{t}$ process with an e^+e^- final state.

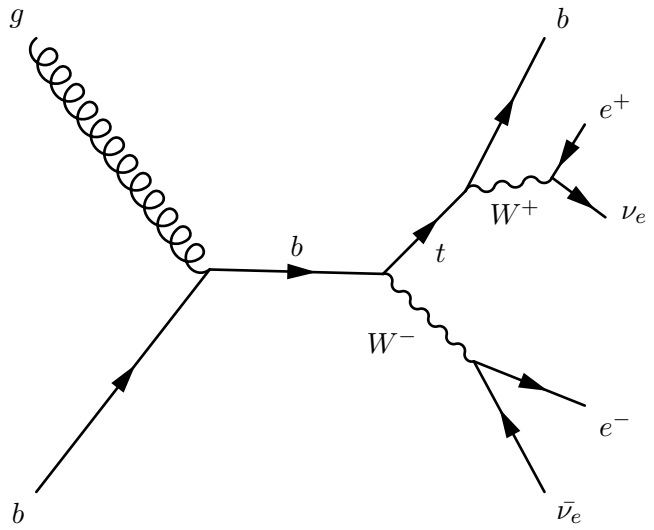


Figure 4.4: Leading order Feynman diagram for Wt process with an e^+e^- final state.

In both analyses a k-factor is derived using the programme Top++2.0 [68] to scale the background from NLO to NNLO in QCD.

Inevitably the un-binned samples lack events at high invariant masses and therefore a fit and extrapolation is used. Two functions are investigated, namely the dijet function: $e^{-a}x^b x^{c\log x}$, and the monomial function: $a/(x+b)^c$. An ensemble of fits is constructed using variations of both fit upper and lower bounds. The best χ^2 fit is used as the nominal estimate. The envelope of all variations is taken as the uncertainty on the background extrapolation.

The $\sqrt{s} = 8$ TeV top sample used MC until 500 GeV then the nominal fit after to give a smooth distribution. Figure 4.5 shows the best fit range is found between the vertical blue lines, the vertical dashed red line denotes the stitching point. The errors on the fitted points are calculated from the envelope of the ensemble of fits. The $\sqrt{s} = 13$ TeV uses a stitching point of 600 GeV, with the MC distribution used below.

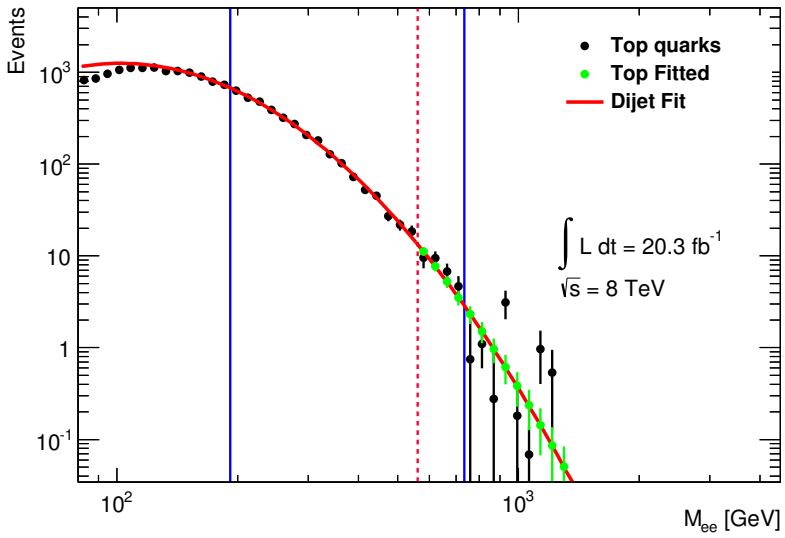


Figure 4.5: $\sqrt{s} = 8$ TeV high invariant mass extrapolation for top background estimate.

4.2.4 Diboson Background

Diboson events (WW, WZ and ZZ) can result in a dielectron final state if the W and Z decay leptonically as shown in Figure 4.6.

The $\sqrt{s} = 8$ TeV analysis samples were generated using HERWIG 6.510 [66] at LO with the CTEQ6L1 PDF [69]. Two mass-binned samples per diboson process are generated, covering true dilepton mass in the range between 400 GeV and 1000 GeV and one above 1000 GeV for a high mass prediction. An un-binned sample for each process is also created to cover the mass range below 400 GeV. The samples are scaled to NLO using a mass-independent k-factor.

The $\sqrt{s} = 13$ TeV analysis generated inclusive samples for each process using SHERPA 2.1.1 [70] at NLO with the CT10 PDF [58]. As only an inclusive sample was generated for each process, a fit and extrapolation is used as described for the top samples. The best fit

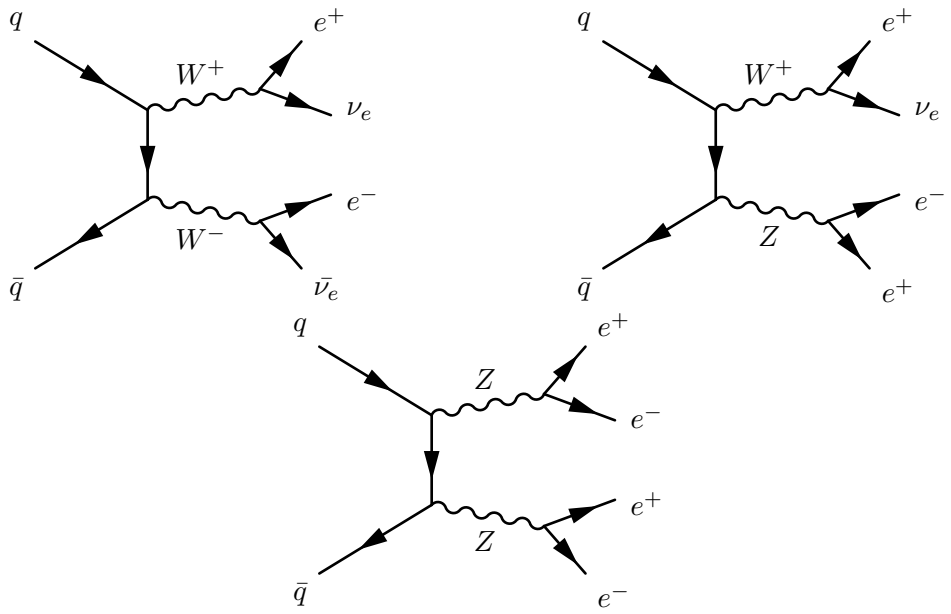


Figure 4.6: Leading-order Feynman diagram for diboson processes with a minimum of a dielectron final state.

is used above a dielectron invariant mass of 280 GeV.

4.2.5 Multi-Jet and W+Jets Background

The final reducible background to be estimated arises from incorrectly identified dielectron events. This contribution consists mainly from the leptonic decay of a W boson and a jet producing an electron-like shower (W+jets), or two jets forming two mistakenly reconstructed electrons (multi-jet). An example leading order diagram for each of these processes is shown in Figure 4.7. Here the method used to estimate this background is discussed, the result of which is used in the kinematic distributions presented in Chapter 5. A full study of the estimate techniques used and derived uncertainties are given in [1] and [4].

The multi-jet and W+jets background is estimated from data using the so-called *matrix* method [71]. Firstly the electron identification and isolation requirement (detailed in Sections 3.2 and 3.3) are loosened in the event selection with regards to the nominal selection. This allows an estimate of the rate at which looser objects pass the nominal selection and and can be used to determine signal contamination by misidentified jets. Two selections are defined as *tight* for the nominal and *loose* for the relaxed criteria,

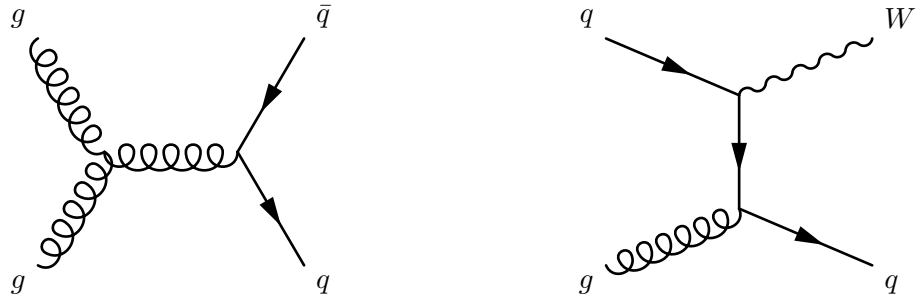


Figure 4.7: Leading order Feynman diagrams for multi-jet and W+jets processes.

where the objects passing the *tight* selection, N_{tight}, are a subset of the *loose* selection, N_{loose}.

A set of four measurable quantities are constructed: N_{TT}, N_{TL} , N_{LT} , N_{LL}, where the first subscript represents the leading E_T object and the second represents the subleading E_T object such that $E_{\mathrm{T,leading}} > E_{\mathrm{T,subleading}}$. A different definition for *loose* is used in these subscripts, where L means an event passes the *loose* selection but fails the *tight* selection, which differs from the basic selection N_{loose} containing all objects in the N_{tight} selection. The subscript T denotes a lepton that passes the *tight* selection. Equation 4.1 shows the relation between these reconstructed quantities and the real ones where the index R represents real electrons and F the objects that have faked electrons;

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} r^2 & rf & fr & f^2 \\ r(1-r) & r(1-f) & f(1-r) & f(1-f) \\ (1-r)r & (1-r)f & (1-f)r & (1-f)f \\ (1-r)^2 & (1-r)(1-f) & (1-f)(1-r) & (1-f)^2 \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} . \quad (4.1)$$

The vector on the right hand side of the equation describes the true quantities such as N_{RR} (the number of events with real leading and subleading electrons) which are not experimentally measurable. The coefficients f and r are called the fake and real efficiencies, respectively. The fake efficiency is the probability that a fake electron passes the loose selection and also the tight selection. The real efficiency is the probability that a real electron passes the loose selection and also the tight selection. Equations 4.2 are used to

determine these efficiencies.

$$f = \frac{N_{tight}^{fake}}{N_{\text{loose}}^{\text{fake}}} \qquad \text{and} \qquad r = \frac{N_{tight}^{real}}{N_{\text{loose}}^{\text{real}}} \ . \tag{4.2}$$

The fake efficiency is determined from fake enriched data samples obtained from events satisfying a number of single electron/electromagnetic object triggers with loosened identification requirements (ranging from 3 – 30% depending on η and E_T). The real efficiency is determined from DY MC (ranging from 90 – 100% depending on η and E_T).

The quantity of interest is the number of events N_{TT} which originate from a pair of objects containing at least one fake ($N_{\text{RR}} = 0$). This is described in Equation 4.3 following the first line of Equation 4.1;

$$N_{\text{TT}}^{\text{Multi-jet and } W\text{+jets}} = rf(N_{\text{RF}} + N_{\text{FR}}) + f^2 N_{\text{FF}} \tag{4.3}$$

and depend on inaccessible truth quantities N_{RF}, N_{FR} and N_{FF}. By inverting the matrix in Equation 4.1 the true variables can be expressed by measurable quantities;

$$\begin{aligned} N_{\text{TT}}^{\text{Multi-jet and } W\text{+jets}} = \frac{1}{(r-f)^2}\big\{&[2rf(f-1)(1-r) + f^2(1-r)^2]N_{\text{TT}} \\ &+ fr^2(1-f)(N_{\text{TL}} + N_{\text{LT}}) \\ &+ f^2 r^2 N_{\text{LL}}\big\} \ . \end{aligned} \tag{4.4}$$

The *matrix* method is not valid in the Z-peak region due to the dominance of real *tight* electrons and can lead to a negative estimate. As such, fits are used in this region. A fit is also used at high mass where the amount of data becomes limited. The dijet and monofit functions are again used with variable ranges to determine the best χ^2 fit as the central value and the remaining envelope of fits used to calculate an error. In the $\sqrt{s} = 8$ TeV ($\sqrt{s} = 13$ TeV) estimate the high mass fit was performed above 500 (600) GeV.

The results of $N_{\text{TT}}^{l+jets \text{ and } di-jets}$ as a function of dielectron mass are used as the fake estimate in the high mass searches and can be seen in Figure 5.1 for the $\sqrt{s} = 8$ TeV analysis and Figure 5.22 for the $\sqrt{s} = 13$ TeV analysis. At $\sqrt{s} = 13$ TeV the number of

events expected at 1 (2) TeV in the $N_{\mathrm{TT}}^{l+jets \text{ and } di-jets}$ mass distribution are 0.14 (0.0007). These are only 3.1 (0.32) % of the DY contribution at the same masses.

4.3 Non-Resonant Signals

4.3.1 ADD Gravitons

To simulate the possible $\sqrt{s} = 8$ TeV ADD signal, 8 separate MC samples using the GRW formalisms are generated with the programme SHERPA 1.3.1 [70] at LO using multi-leg matrix elements and the CTEQ6L1 PDF [69]. The generation includes the production of up to one additional jet from QCD initial state radiation. These are defined at different ultraviolet cutoff scales, M_{S}, of 2.5, 3, 3.25, 3.5, 3.75, 4, 4.75 and 50 TeV. These samples include contributions from SM DY and the ADD Gravitons. To ensure a large number of events are generated at high mass, three mass binned samples are created covering the dielectron invariant mass range 300 – 600 GeV, 600 – 1200 GeV and >1200 GeV. With a $M_{\mathrm{S}} = 50$ TeV sample, the scale of new physics is much beyond the reach of this analysis, and as such a DY only SHERPA sample can be obtained in the search region. This SHERPA DY sample can then be removed from all other signals to leave a LO ADD signal only contribution. The results obtained with the GRW formalism can be converted into the Hewett and HLZ ($n > 2$) formalisms and do not need dedicated generation. The HLZ $n = 2$ requires a dedicated sample due to the extra dependence on the dilepton mass (Equation 1.5).

4.3.2 Contact Interactions

To simulate the possible $\sqrt{s} = 13$ TeV CI signal, a re-weighting algorithm is applied to the LO DY samples generated using the programme PYTHIA8 [59] NNPDF23LO [72]. The algorithm takes as input on an event-by-event basis: the desired Λ value and model for CI, as well as the incoming/outgoing particle four-momentas, and the incoming quark flavour. The LO cross-section for SM DY and the CI model of choice are calculated for each event and the ratio of the two is applied to the event as a weight to scale the SM DY spectrum into the desired CI models spectrum.

So that the signal samples are estimated to the same order as the main NNLO DY

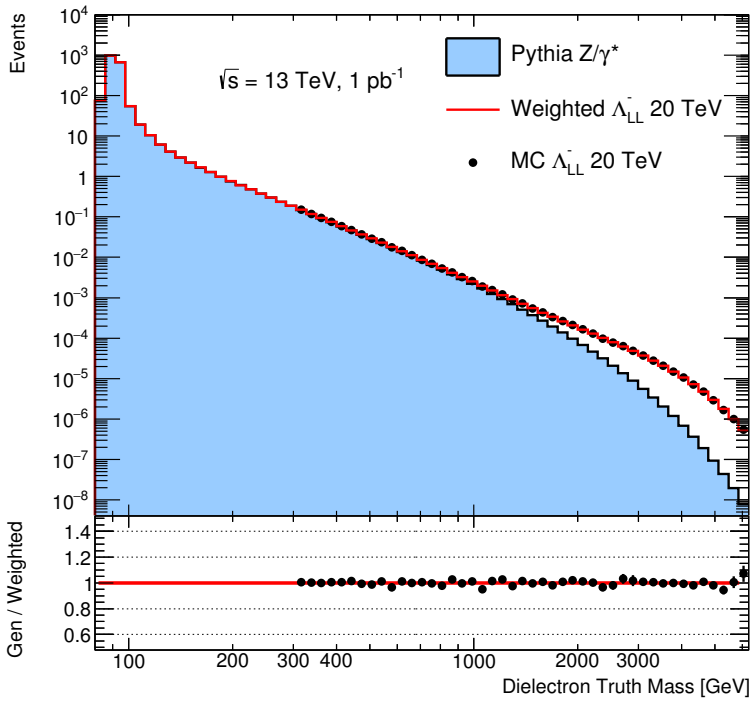


Figure 4.8: Generator level LL CI $\Lambda = 20$ TeV constructive model dielectron invariant mass spectrum re-weighted from LO DY.

background estimate, a k-factor is calculated using FEWZ to scale the LO DY sample used for the CI re-weighting to NNLO in QCD and NLO in EW.

To validate this technique, dedicated CI samples were generated using the programme PYTHIA8 [59] with the NNPDF23LO [72] PDF. To ensure enough events are generated at high mass, three mass binned samples are created in four mass bins covering the dielectron invariant mass range 300 – 600 GeV, 600 – 1200 GeV, 1200 – 3000 GeV and >3000 GeV. In the region below 300 GeV a negligible amount of new physics signal is expected, therefore the SM DY estimate is used. The validation of a few signal re-weightings produced at truth-level are presented in Figures 4.8, 4.9, 4.10 and 4.11. The re-weighting procedure is seen to accurately replicate the dedicated signal MC sample, with the slight variations occurring due to a lack of events at the highest masses in each binned sample.

The CI re-weighted samples are combined estimates from both the CI signal and DY background processes. The DY sample used in the re-weighting is subtracted from the CI re-weighted samples to produce a pure CI signal estimate.

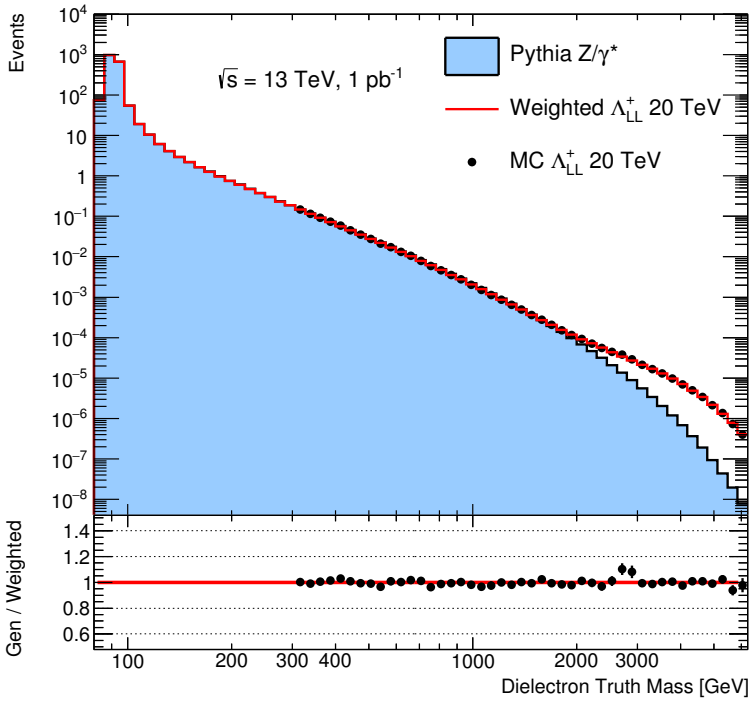


Figure 4.9: Generator level LL CI $\Lambda = 20$ TeV destructive model dielectron invariant mass spectrum re-weighted from LO DY.

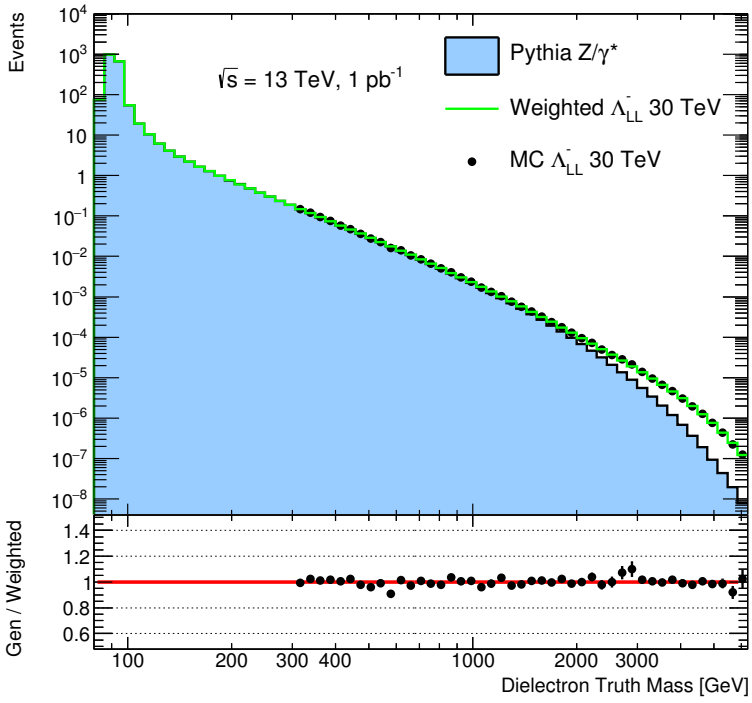


Figure 4.10: Generator level LL CI $\Lambda = 30$ TeV constructive model dielectron invariant mass spectrum re-weighted from LO DY.

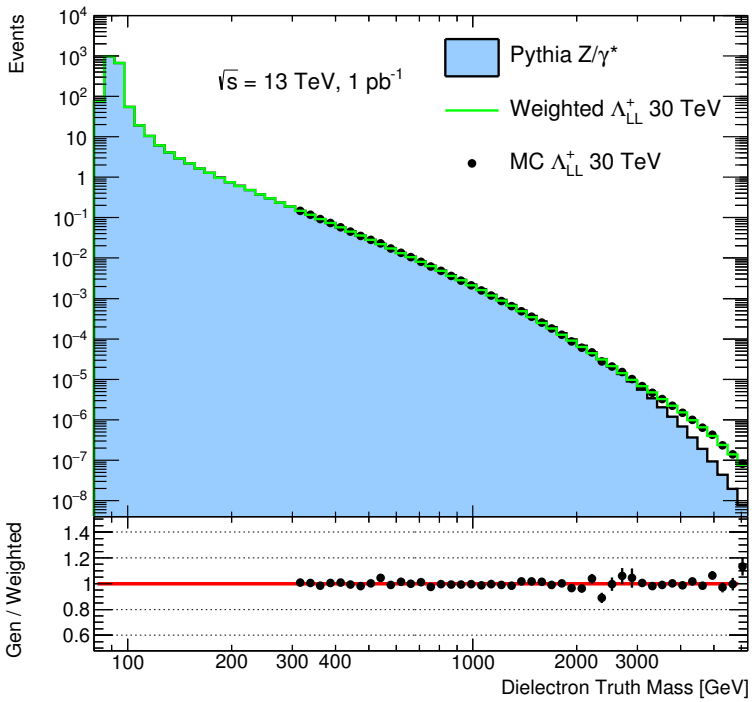


Figure 4.11: Generator level LL CI $\Lambda = 30$ TeV destructive model dielectron invariant mass spectrum re-weighted from LO DY.

"Not everything that counts can be counted, and not everything that can be counted counts."

–William Bruce Cameron

5

Event Selection

In order to search for new physics with a dielectron final state, signal events have to be efficiently selected and background rejected. This chapter describes the selection of events with dielectron candidates. The same selection is conducted on data as for the background and signal estimation. The outline of the $\sqrt{s} = 8$ TeV analysis selection is discussed first, followed by the $\sqrt{s} = 13$ TeV analysis, including the signal efficiency checks of selection criteria. The relative and cumulative efficiencies of the selection for the search regions are displayed for the irreducible DY background. Finally a data-background comparison is presented with a selection of possible signals overlaid. Included in the kinematic variable comparisons are systematic uncertainties displayed as bands in the inset ratio, the sources of which are discussed in Chapter 6. In both searches an invariant mass normalisation region is defined as 80–120 GeV in order to scale MC to data in the Z-peak region.

5.1 ADD Analysis

The $\sqrt{s} = 8$ TeV analysis selects the two highest E_T, isolated, well reconstructed electrons in an event and uses the invariant mass as a discriminating variable to search for an ADD LED signature. The ordering of the event selection is detailed below.

Event level selection:

- Event is required to have fired a two electromagnetic object trigger. The E_T of the leading (subleading) object has to be > 35 GeV (25 GeV). Both objects must pass the *loose* shower shape identification variables. In ATLAS this is known as the EF_g35_loose_g25_loose trigger.
- Event contains a reconstructed primary vertex[1] with at least two tracks.
- Event contains at least two electrons.

Electron level selection:

- Electron is within $|\eta| < 2.47$, excluding the calorimeter crack region $1.37 < |\eta| < 1.52$.
- Electron not flagged as being from a bad calorimeter cluster.
- Subleading electron must have $E_T > 30$ GeV, and the leading electron $E_T > 40$ GeV.
- Electron passes *medium++* identification criteria.

Dielectron pair selection:

- Select two highest E_T electrons.
- Leading electron must be isolated with: $E_T Cone20 < 0.007 E_T + 5.0$ GeV.
- Subleading electron must be isolated with: $E_T Cone20 < 0.022 E_T + 6.0$ GeV.
- Leading and subleading electron must have opposite-sign charge.
- Dielectron invariant mass > 80 GeV.

5.1.1 Acceptance × Efficiency

The detector acceptance × selection efficiency of each criterion is presented in Table 5.1 for $m_{ee} > 1000$ GeV where the search region optimisation begins (Section 7.4.1). The DY MC is used to calculate the acceptance × efficiency of signal as it is irreducible and because it contains a greater number of events than the dedicated signal samples. The

[1] The vertex with the largest $\Sigma {p_T}^2$ of the associated tracks.

introduction of an opposite sign charge selection is due to the use of angular information in the $\sqrt{s} = 8$ TeV CI search, where the charge orientation of the lepton pair needs to be known. Whilst this decreases efficiency and adds an additional systematic uncertainty due to charge mis-identification, the loss is slightly compensated by the increase in multi-jet and W+jets rejection. A study showed the ADD expected limits dropped by $\sim 1 - 2\%$.

Table 5.1: Dielectron channel cutflow table, presenting the relative and cumulative efficiencies for each important selection criterion. These values are given for the irreducible DY background at dielectron masses greater than 1000 GeV where signal could be present. Cuts with $\sim 100\%$ efficiency are not included.

Criterion	DY ($m_{ee} > 1000$ GeV)	
	Rel. [%]	Cum. [%]
Trigger	92.70 ± 0.03	92.70 ± 0.03
η	97.48 ± 0.02	90.28 ± 0.03
E_T	94.33 ± 0.03	85.02 ± 0.04
ID	90.03 ± 0.03	76.54 ± 0.05
Leading Isolation	98.96 ± 0.01	75.75 ± 0.05
Subleading Isolation	97.93 ± 0.02	74.18 ± 0.05
Opposite-Sign	89.11 ± 0.04	66.11 ± 0.06

5.1.2 Data-Background Comparisons

The data and total background estimate at $\sqrt{s} = 8$ TeV is presented for the invariant mass spectrum with possible ADD signals overlaid in Figure 5.1. Other kinematics distributions of the leading and subleading E_T, η, and ϕ are found in Figures 5.2, 5.3 and 5.4, respectively. Additional data-background comparisons, splitting the distributions into leading and subleading electrons separately, are provided in Appendix A. The electron E_T shows some disagreement between data and expected events. This disagreement is understood to be mainly due to higher-order QCD effects (additional jets) in the Z region; a study of this is shown in Section 5.2.4. A study investigating the impact of re-weighting the DY MC to match the full dielectron p_T spectrum in data showed a negligible impact on the dilepton invariant mass distribution. These studies are detailed in Appendix B.

The expected and observed number of events with the 20.3 fb^{-1} dataset are presented in Table 5.2 for the ADD search region. The ADD search region is defined from 1900 – 4500 GeV; this choice is motivated by an optimisation performed using the highest expected limit as described in Section 7.4.1.

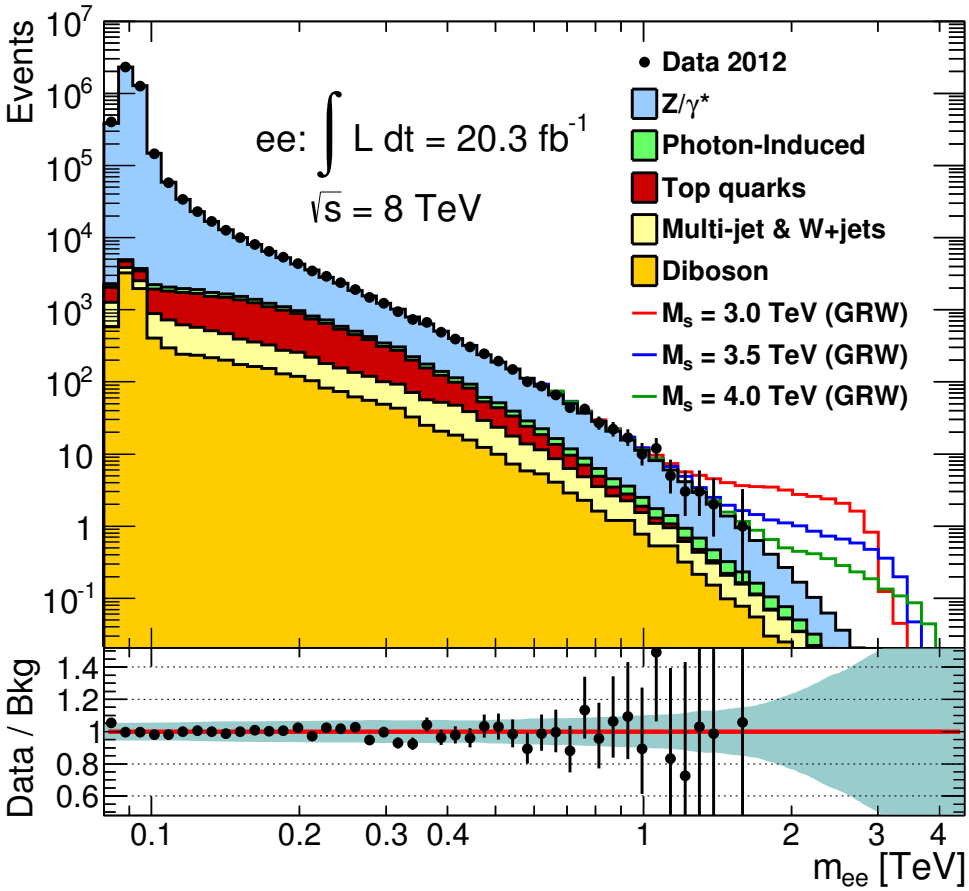


Figure 5.1: $\sqrt{s} = 8$ TeV dielectron invariant mass spectrum showing data and background processes with three ADD GRW signals overlaid, the bottom inset shows the ratio between data and background including the total systematic variation displayed as a band.

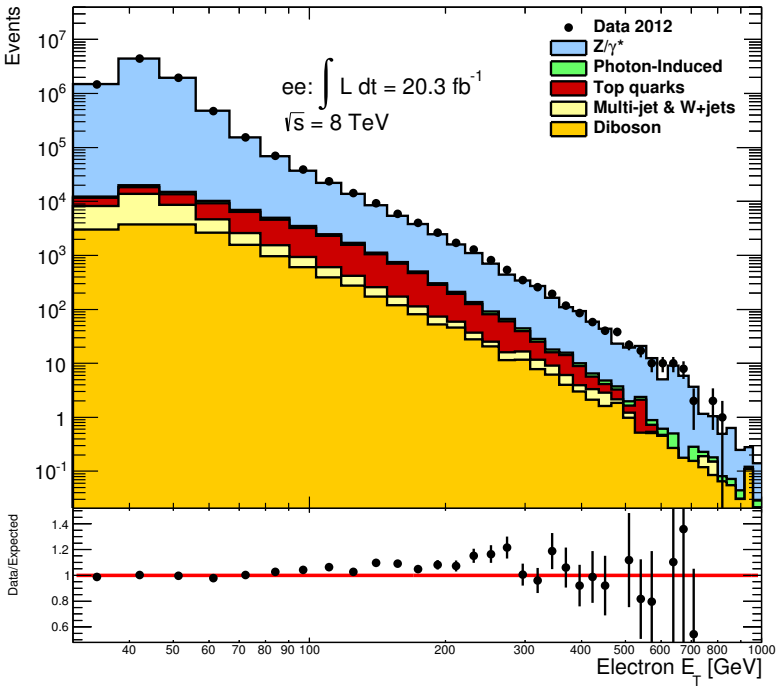


Figure 5.2: $\sqrt{s} = 8$ TeV E_T distribution for both leading and subleading electrons.

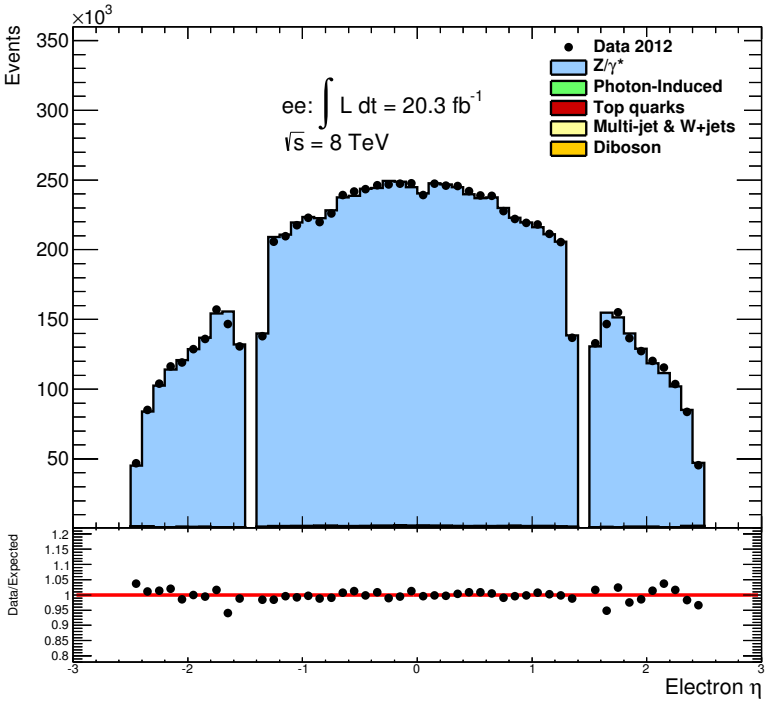


Figure 5.3: $\sqrt{s} = 8$ TeV η distribution for both leading and subleading electrons.

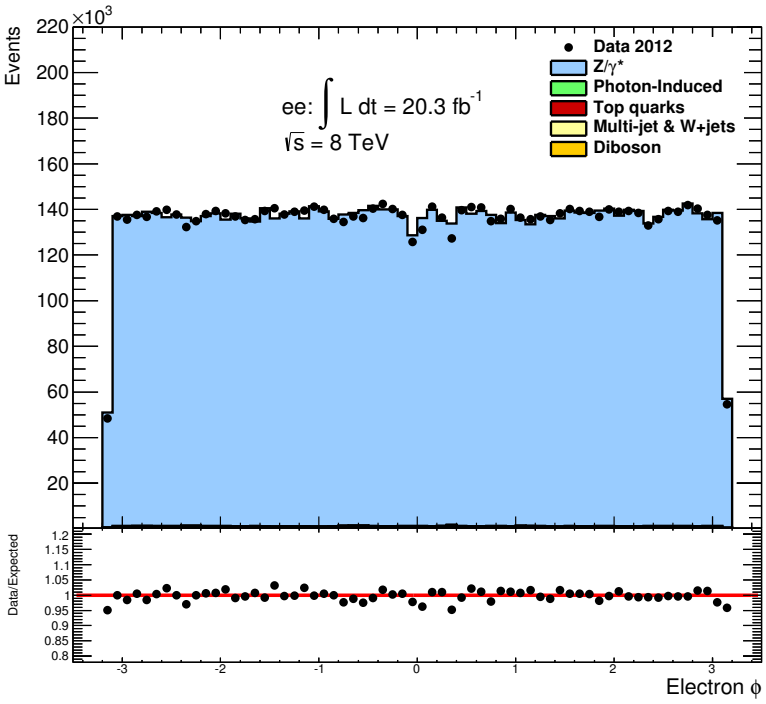


Figure 5.4: $\sqrt{s} = 8$ TeV ϕ distribution for both leading and subleading electrons.

Table 5.2: Table presenting the expected and observed number of events in the signal region of the dielectron channel search for ADD LED. Expected yields are provided for benchmark ADD signal models under the GRW formalism. The errors quoted originate from both MC statistical and systematic uncertainties.

Process	$1900 \leq m_{ee} \leq 4500$ GeV
Drell–Yan	0.43 ± 0.12
Top quarks	< 0.002
Multi-jet & W+jets	0.062 ± 0.012
Diboson	0.053 ± 0.005
Photon-Induced	0.06 ± 0.06
Total SM	0.61 ± 0.13
Data	0
SM+ADD ($M_S = 3.25$ TeV)	9.7 ± 0.9
SM+ADD ($M_S = 3.50$ TeV)	5.8 ± 0.5
SM+ADD ($M_S = 3.75$ TeV)	3.73 ± 0.34
SM+ADD ($M_S = 4.00$ TeV)	2.56 ± 0.24
SM+ADD ($M_S = 4.75$ TeV)	1.21 ± 0.15

5.2 Contact Interaction Analysis

As the LHC reaches closer to the design centre-of-mass energy the relative cross-section increase from $\sqrt{s} = 8$ TeV for SM processes is typically 2-4 times higher, while many BSM processes increase by factors of 4 upwards [73, 74]. To maximise the search efficiency at these new sensitivities, the $\sqrt{s} = 13$ TeV analysis is an evolution of the $\sqrt{s} = 8$ TeV analysis. The ordering of the event selection is detailed below.

Event level selection:

- Event is required to have fired a two electron trigger. The E_T of each electron has to be > 17 GeV. Both electrons must pass the likelihood *loose* electron identification. In ATLAS this is known as the 2e17_lhloose trigger.
- Event contains at least two electrons.

Electron level selection:

- Electron is within $|\eta| < 2.47$, excluding the calorimeter crack region $1.37 < |\eta| < 1.52$.
- Electron not flagged as being from a bad calorimeter cluster.
- Electron must have $E_T > 30$ GeV.
- Electron passes the inner detector track requirement: $|\sigma_{d_0^{BL}}| < 5$.
- Electron must pass the likelihood *medium* identification criteria.
- Electron passes *loose* isolation criteria.

Dielectron pair selection:

- Select two highest E_T electrons.
- Dielectron invariant mass $>$ 80 GeV.

5.2.1 Track-Vertex Association Efficiency

To select electrons from the hard scatter process, the ATLAS inner tracking group made recommendations for run 2 on electron track impact parameter requirements. The main variable of interest is the transverse impact parameter d_0^{BL}: this is the distance of closest

approach of a track to the beam line. Figure 5.5 shows how charged tracks from a primary vertex will have a different distance of closest approach than tracks from a secondary vertex (B meson decays for instance) which will have travelled some distance before decaying (typically 0.5 mm [5]). By extrapolating the tracks of charged particles the position of closest approach can be determined, depending on the size of the transverse distance from the beam line, the particle can be associated to a secondary decay vertex.

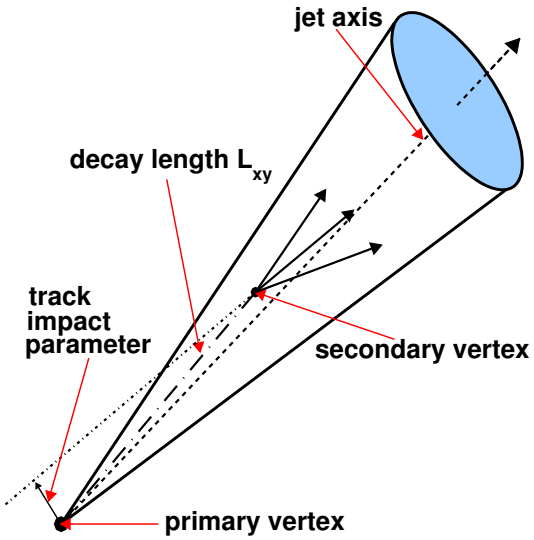


Figure 5.5: Track impact parameter d_0 used in heavy flavour decay discrimination [75].

The ATLAS recommended threshold for electrons is to use a d_0 significance (defined in Table 3.2) of $|\sigma_{d_0^{BL}}| < 5$. This choice is motivated to be loose enough such that there is minimal signal acceptance loss for ATLAS analyses selecting electrons and determined from the Z mass region measurements. As this is a high mass search, a check of this was performed to investigate the loss of signal efficiency and potential background rejection with the introduction of this cut.

The data-background d_0 significance distributions applying the nominal event selections except the d_0 significance requirement are presented in Figures 5.6 and 5.7 for the leading and subleading electrons, respectively. This shows the dominance of the irreducible DY throughout the spectrum, however, a difference in shape could result in a some reducible background rejection. Normalising the d_0 significance distributions for the MC backgrounds, it can be shown that at higher $|\sigma_{d_0^{BL}}|$ values the reducible backgrounds

dominate over the signal-like acceptance of DY (Figure 5.8).

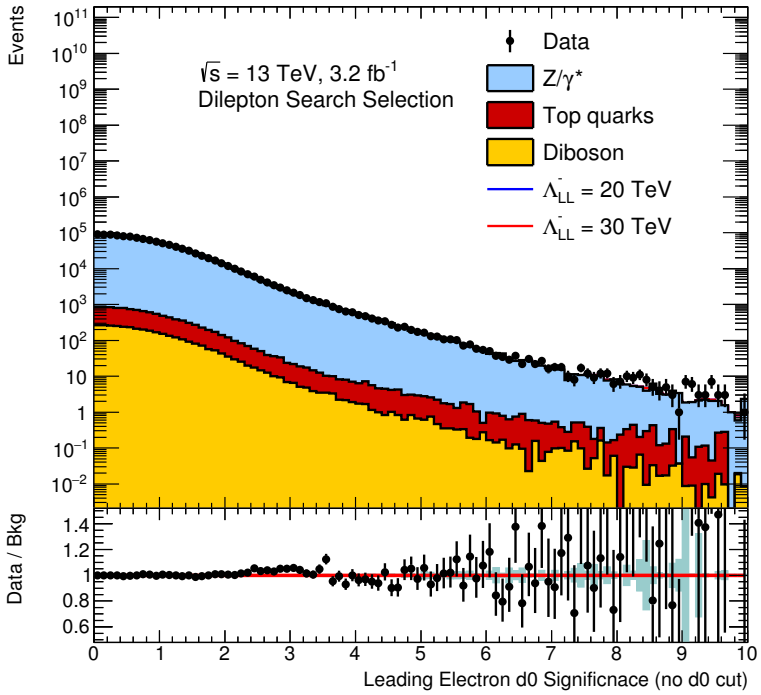


Figure 5.6: Leading electron d_0 significance distributions after the dielectron search event selection (except d_0 criterion).

The efficiency of including this criteria can be seen to have minimal impact on the signal-like DY efficiency and an increased rejection of real top contributions from heavy flavour decays (Figures 5.9 and 5.10 for leading and subleading electron E_T and 5.11 and 5.12 for leading and subleading electron η). The same effect is seen as a function of invariant mass (Figure 5.13), the signal-like DY efficiency remains extremely high, while there is a slight rejection of backgrounds from top processes. The choice to remain at the recommended selection value of $|\sigma_{\mathrm{d_0^{BL}}}| < 5$ is motivated by the near 100% signal efficiency at high dielectron invariant mass.

A secondary cut was investigated on the longitudinal impact parameter with respect to the vertex of the hard scatter interaction, z_0. While the introduction of this cut can reduce background from cosmic rays in the muon final state search, it would not remove any additional background from the electron final state as the contribution from cosmic rays is already negligible. A study showed a few percent loss at low E_T in signal efficiency when using this criterion and as such was not adopted into the nominal event selection.

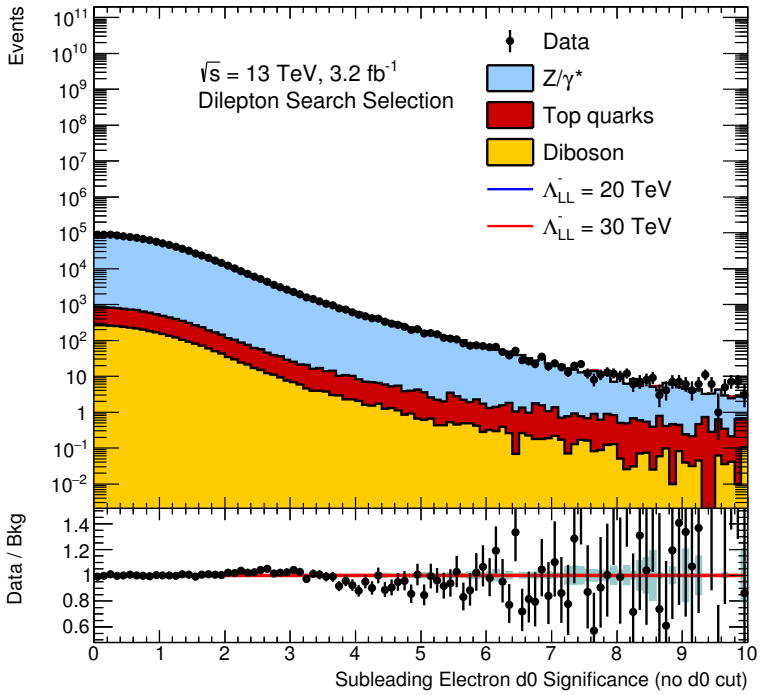


Figure 5.7: Subleading electron d_0 significance distributions after the dielectron search event selection (except d_0 criterion).

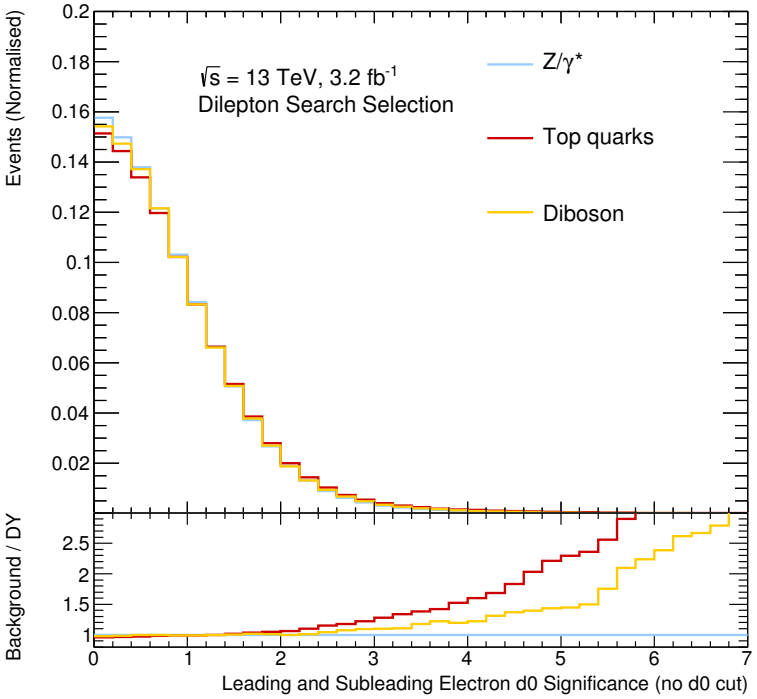


Figure 5.8: Leading and subleading electron d_0 significance distributions after the dielectron search event selection (except d_0 criterion), normalised to unit area.

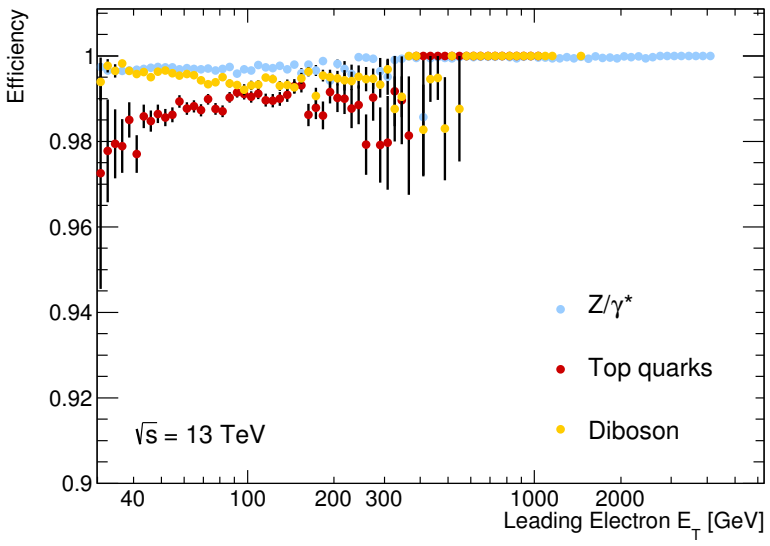


Figure 5.9: Lead electron d_0 significance criterion efficiency versus E_T.

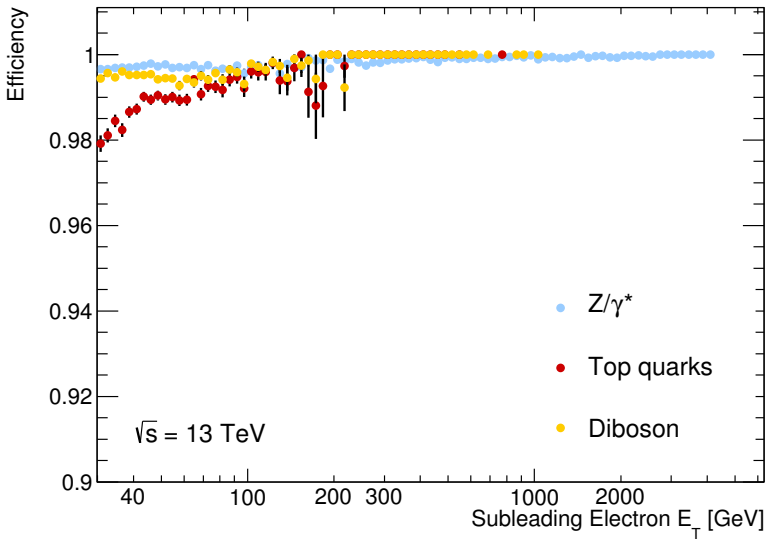


Figure 5.10: Subleading electron d_0 significance criterion efficiency versus E_T.

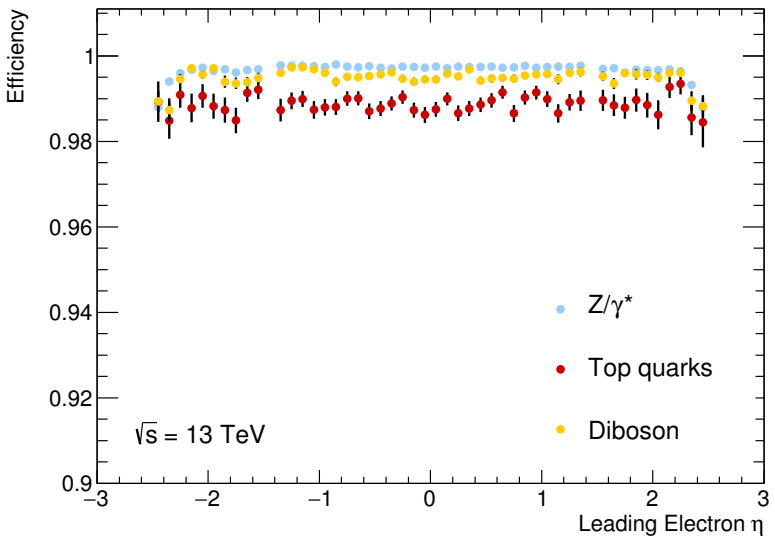


Figure 5.11: Leading electron d_0 significance criterion efficiency versus η.

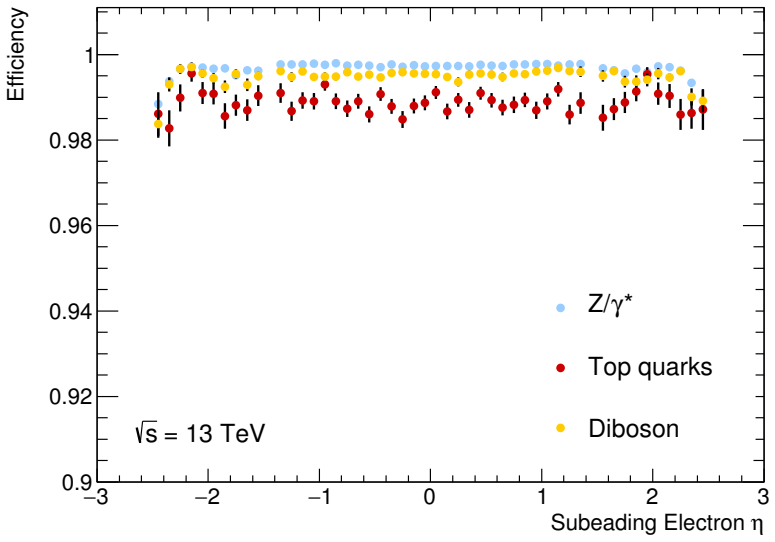


Figure 5.12: Subleading electron d_0 significance criterion efficiency versus η.

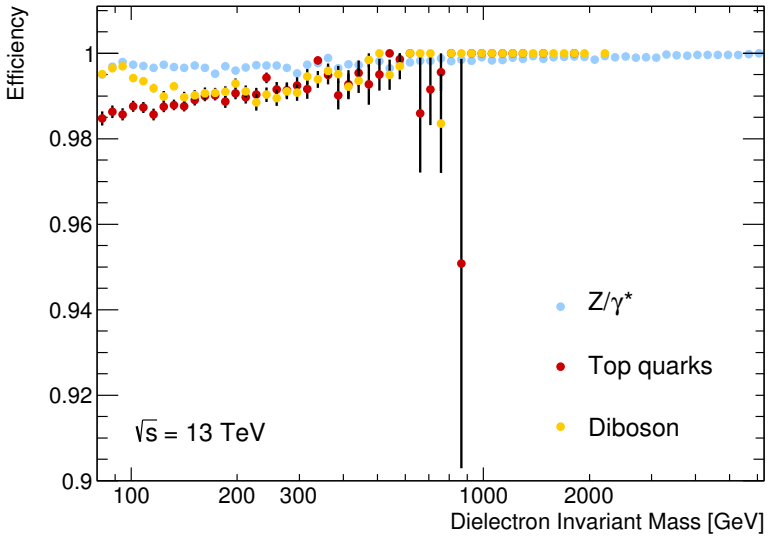


Figure 5.13: Electron d_0 significance criterion efficiency versus the reconstructed dielectron invariant mass.

5.2.2 Electron Identification Efficiency

In order to choose an electron identification criteria, a study was performed to compare the different likelihood-based working points, namely: *loose*, *medium*, and *tight*. In run 2, these working point efficiencies were determined by the ATLAS electron/photon performance group from $Z \to ee$ and di-jet simulation, with data derived corrections to shower shape variables.

The DY MC is used to test the signal-like efficiency versus leading and subleading electron E_T, and η. The results of this study are shown in Figures 5.14, 5.15, 5.16 and 5.17 for different working points using a minimal preselection of $E_T > 30$ GeV and the nominal selection η acceptance.

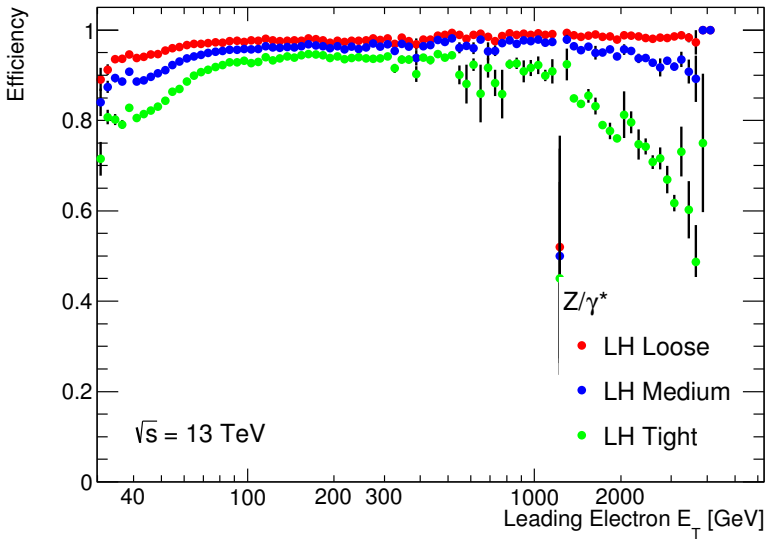


Figure 5.14: Electron ID efficiency for various working points, as applied to DY MC to represent the expected signal efficiency versus leading electron E_T.

The real electron component of the reducible background efficiency for tops and dibosons, and the DY process is also presented in Figures 5.18, 5.19 and 5.20 versus dilepton truth mass for the different working points. This study shows that likelihoods (LH) *loose* and *medium* provide a similar performance in terms of signal efficiency compared to the background rejection for the dielectron search. The *tight* criteria displays a loss of signal acceptance at very high E_T, as such this working point was ruled out for use in the $\sqrt{s} = 13$ TeV search. The fake efficiency is extremely low using the electron LH identi-

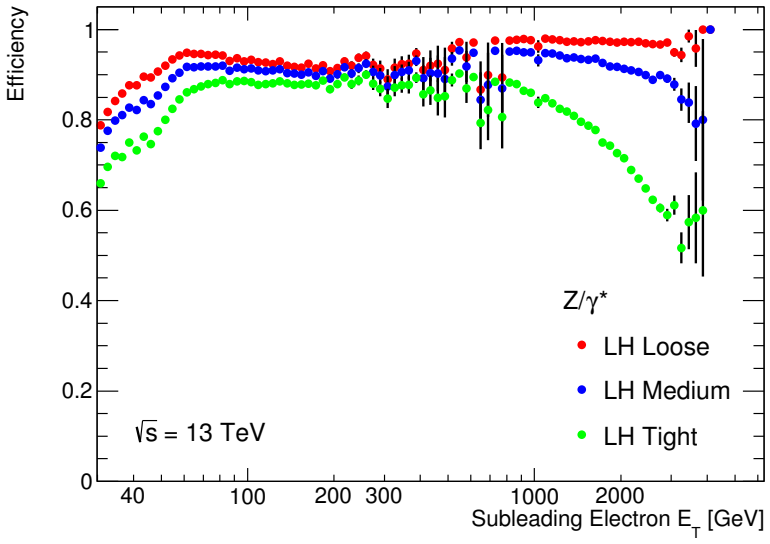


Figure 5.15: Electron ID efficiency for various working points, as applied to DY MC to represent the expected signal efficiency versus subleading electron E_T.

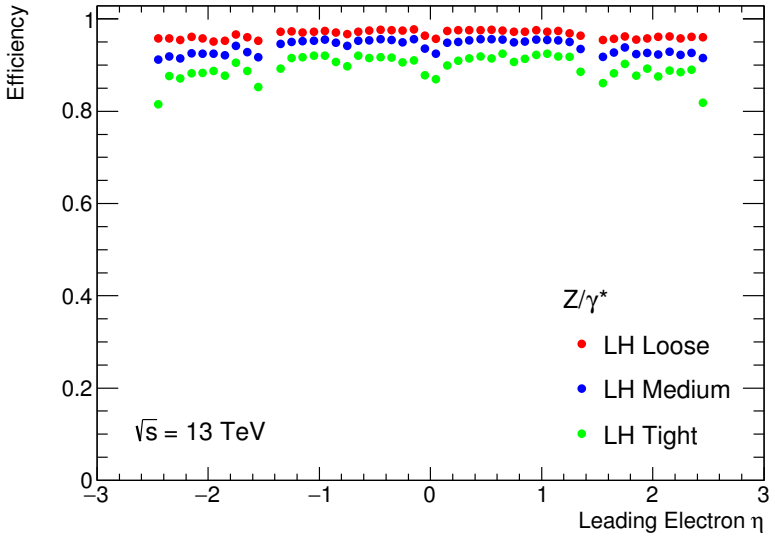


Figure 5.16: Electron ID efficiency for various working points, as applied to DY MC to represent the expected signal efficiency versus leading electron η.

fication. At 3 TeV the *medium* (*loose*) working point rejects all but 0.1 (0.2) % of the background [4].

The *medium* working point is ultimately chosen for the nominal dielectron event selection, as it makes the analysis more robust against fake electrons, as well as being important

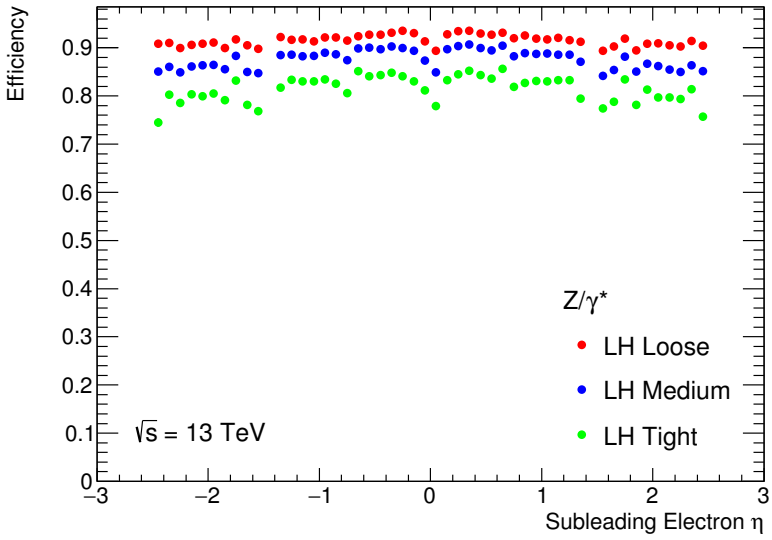


Figure 5.17: Electron ID efficiency for various working points, as applied to DY MC to represent the expected signal efficiency versus subleading electron η.

for the data-driven background estimate which requires a second looser selection. Each of the *medium* working point efficiencies for the three processes used in this study are shown in Figure 5.21 as a function of dilepton invariant truth mass.

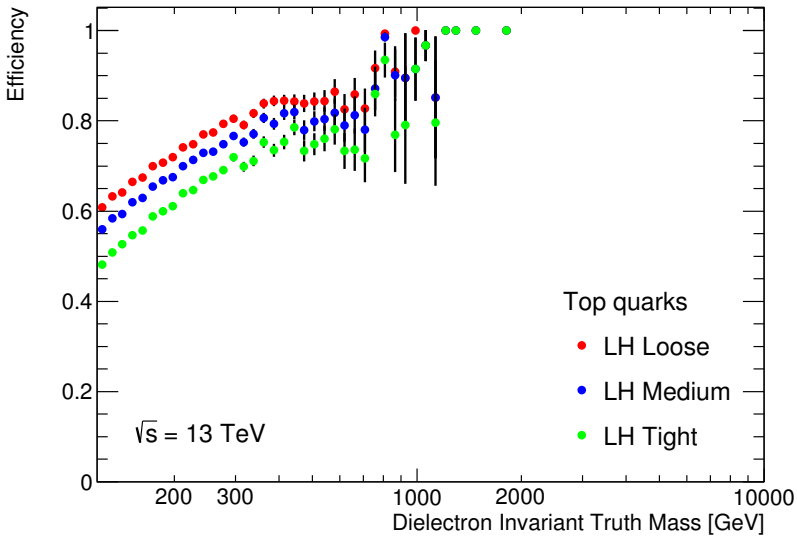


Figure 5.18: Electron ID efficiency for various working points as applied to the real contributions from top quark processes versus dilepton truth mass.

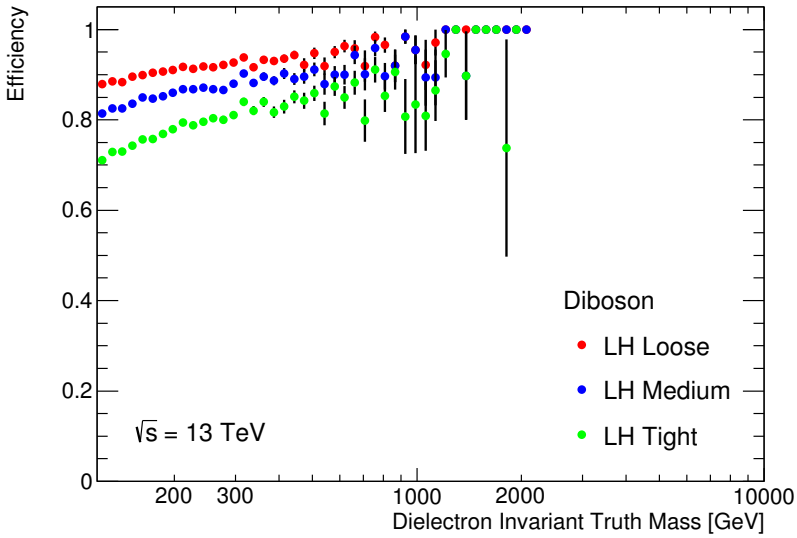


Figure 5.19: Electron ID efficiency for various working points as applied to the real contributions from diboson processes versus dilepton truth mass.

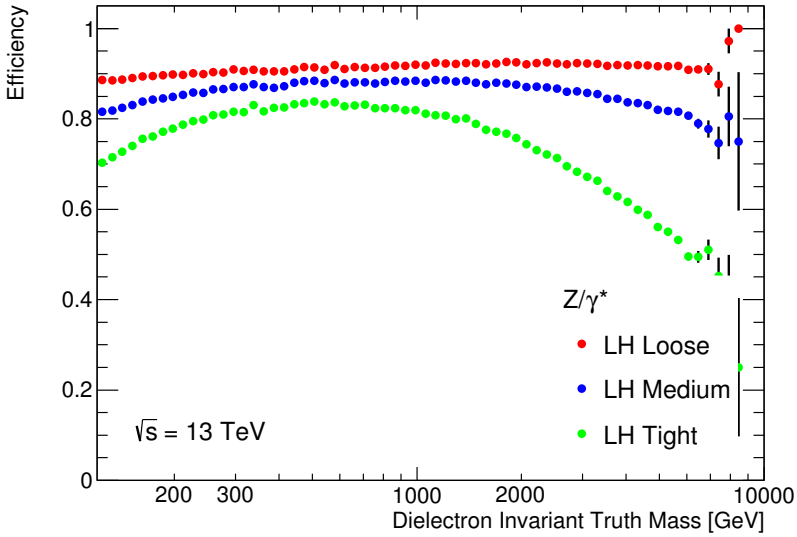


Figure 5.20: Electron ID efficiency for various working points as applied to the real contributions from DY processes versus dilepton truth mass.

5.2.3 Acceptance × Efficiency

The detector acceptance × selection efficiency of each selection is presented in Table 5.3 for $m_{ee} > 400$ GeV where the signal search region begins. DY MC is used instead of signal due to the same expected acceptance but with a greater number of sample events.

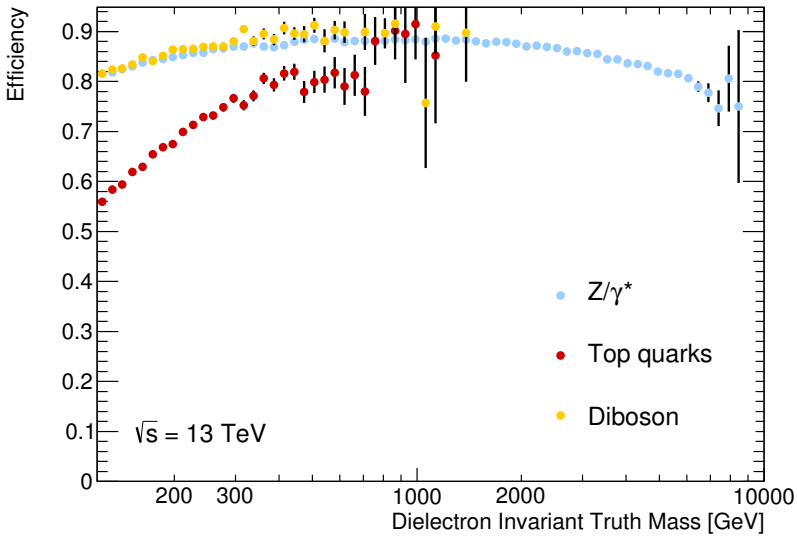


Figure 5.21: The electron ID efficiencies for the Medium working point as applied to DY, and real components of tops, and diboson processes versus dilepton truth mass.

Table 5.3: Dielectron channel cutflow table, presenting the relative and cumulative efficiencies for each important selection criterion. These values are given for the dominant DY background at dielectron masses greater than 400 GeV where signal could be present. Cuts with $\sim 100\%$ efficiency are not included.

Criterion	DY ($m_{ee} > 400$GeV)	
	Rel. [%]	Cum. [%]
Trigger	80.03 ± 0.05	80.03 ± 0.05
η	97.18 ± 0.02	77.75 ± 0.05
E_T	94.31 ± 0.03	73.20 ± 0.05
d_0	99.67 ± 0.01	72.96 ± 0.05
ID	92.74 ± 0.03	67.66 ± 0.06
Isolation	98.53 ± 0.01	66.66 ± 0.06

5.2.4 Data-Background Comparisons

After event selection the data and total background estimate is presented for the invariant mass spectrum with possible non-resonant CI signals overlaid in Figure 5.22.

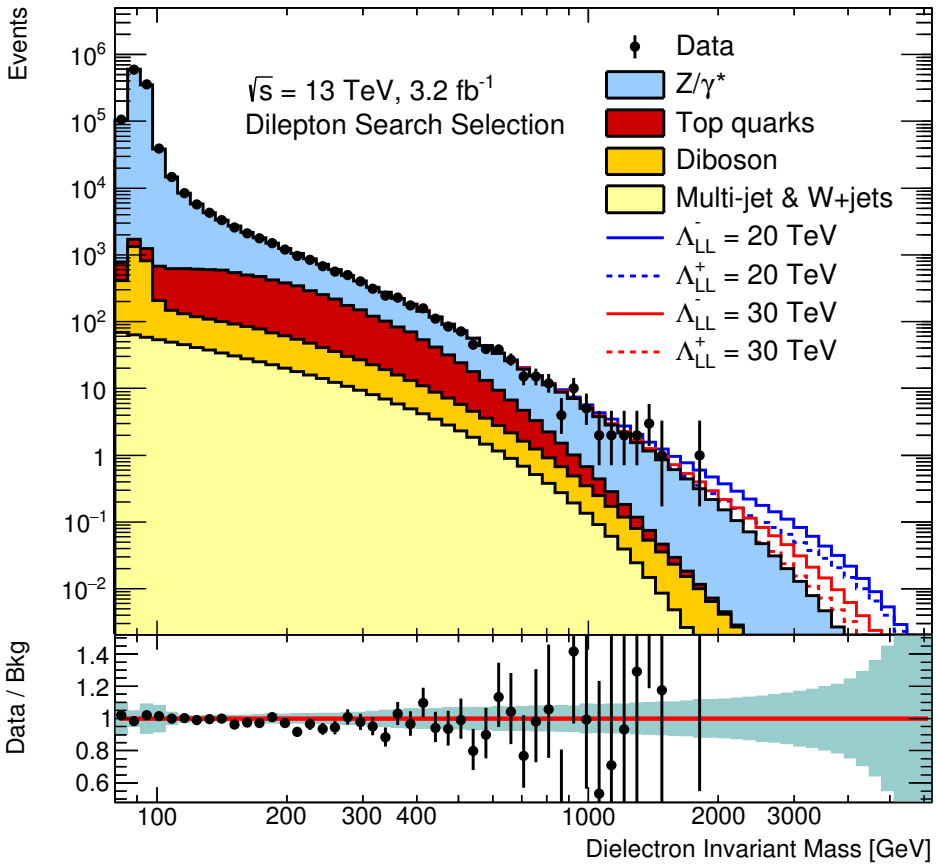


Figure 5.22: Invariant mass spectrum of data and background processes with CI signal overlaid, the bottom inset shows the ratio between data and background.

The highest invariant mass event is displayed using the ATLAS Atlantis programme in Figure 5.23. The leading electron has an E_T of 373 GeV and an η of -1.03. The subleading electron has an E_T of 246 GeV and an η of 2.45. The invariant mass of the pair is 1775 GeV.

Other kinematics distributions of the leading and subleading E_T, η, and ϕ are found in Figures 5.24, 5.25 and 5.26. Additional data-background comparisons, splitting the distributions into those coming from the leading and subleading electrons separately, are provided in Appendix A.

The leading and subleading electron E_T distribution (Figure 5.24) shows some disagreement between data and expected events, as in the $\sqrt{s} = 8$ TeV analysis (Figure 5.2). The level of disagreement is reduced if the distributions are plotted for events above the Z-region ($m_{ee} > 120$ GeV), as shown in Figure 5.27. This is due to the mis-modelling of higher-order QCD effects in the Z-region [76]. The corresponding kinematically constrained ($m_{ee} > 120$ GeV) η and ϕ distributions are shown in Figures 5.28 and 5.29, respectively.

The summary of expected and observed number of events with the 3.2 fb^{-1} dataset are presented in Table 5.4 for large mass windows in the CI search region. The search region is defined from 400 – 6000 GeV in 8 increasing bin sizes. The expected and observed number of events can be seen for each of these bins in Figure 5.30.

Table 5.4: Expected and observed event yields in the dielectron channel in different dilepton mass intervals. The quoted errors correspond to the combined statistical and systematic uncertainties.

m_{ee} [GeV]	500–700	700–900	900–1200	1200–1800	1800–3000	3000–6000
Drell–Yan	149 ± 10	38.3 ± 3.0	16.5 ± 1.4	5.6 ± 0.6	0.78 ± 0.10	0.030 ± 0.005
Top quarks	47.2 ± 1.6	6.2 ± 0.8	1.13 ± 0.35	0.12 ± 0.09	0.002 ± 0.006	<0.001
Diboson	10.3 ± 0.9	2.3 ± 0.5	0.78 ± 0.28	0.20 ± 0.11	0.021 ± 0.018	<0.001
Multi-jet & W+jets	7.2 ± 1.8	1.6 ± 0.8	0.5 ± 0.4	0.08 ± 0.10	0.002 ± 0.005	<0.001
Total SM	214 ± 11	48.4 ± 3.2	18.9 ± 1.6	6.0 ± 0.6	0.81 ± 0.10	0.030 ± 0.006
Data	202	44	17	9	0	0
SM+CI (Λ^{-}_{LL} = 20 TeV)	220 ± 11	52.1 ± 3.2	22.2 ± 1.6	8.8 ± 0.6	2.22 ± 0.14	0.289 ± 0.018
SM+CI (Λ^{+}_{LL} = 20 TeV)	210 ± 11	47.0 ± 3.2	18.4 ± 1.6	6.2 ± 0.6	1.29 ± 0.10	0.187 ± 0.012
SM+CI (Λ^{-}_{LL} = 30 TeV)	217 ± 11	49.9 ± 3.2	20.2 ± 1.6	7.0 ± 0.6	1.23 ± 0.10	0.094 ± 0.007
SM+CI (Λ^{+}_{LL} = 30 TeV)	212 ± 11	47.6 ± 3.2	18.5 ± 1.6	5.8 ± 0.6	0.81 ± 0.10	0.049 ± 0.006

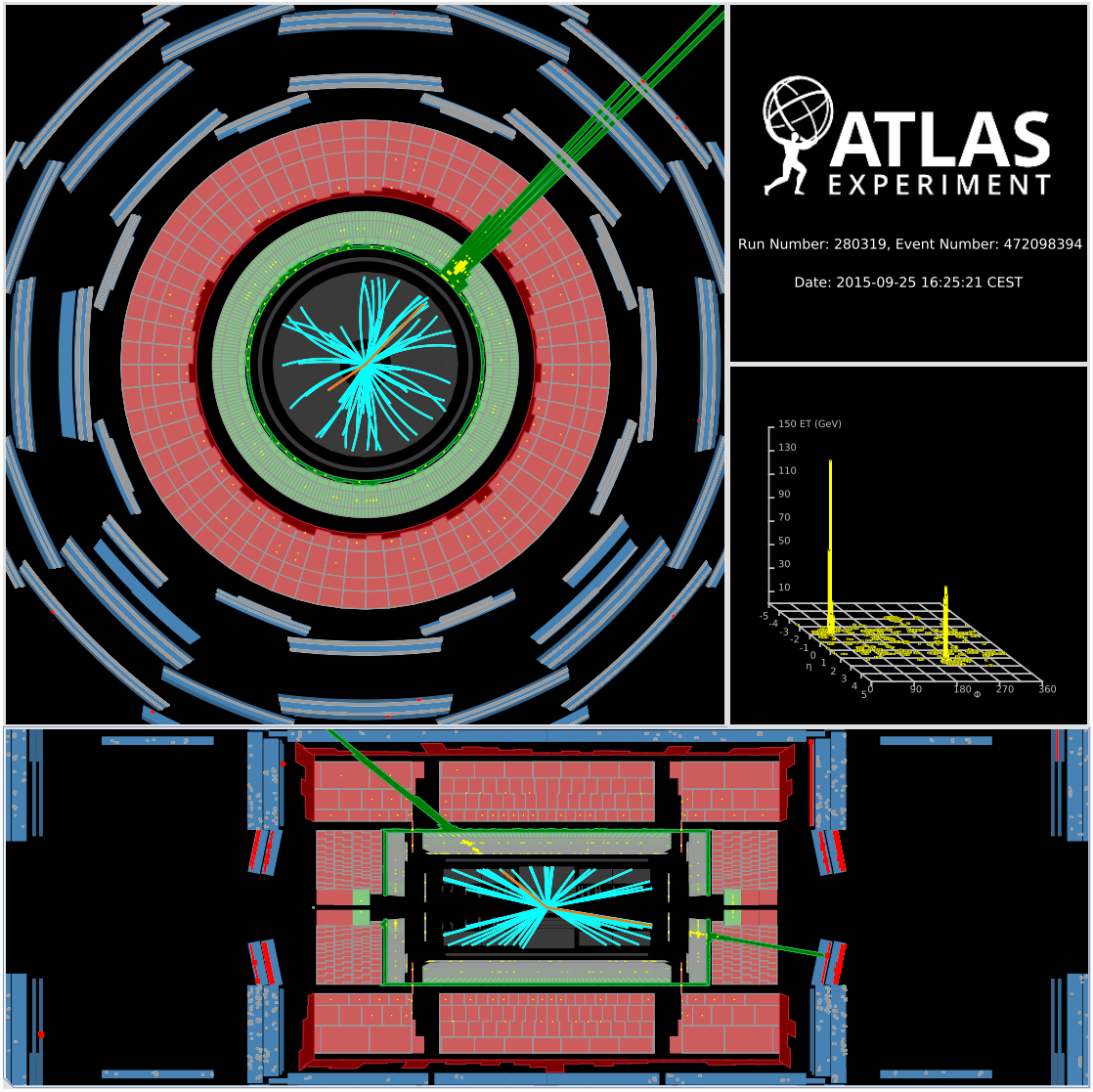


Figure 5.23: Event display of highest invariant mass dielectron event (1775 GeV) in the $\sqrt{s}$ = 13 TeV analysis.

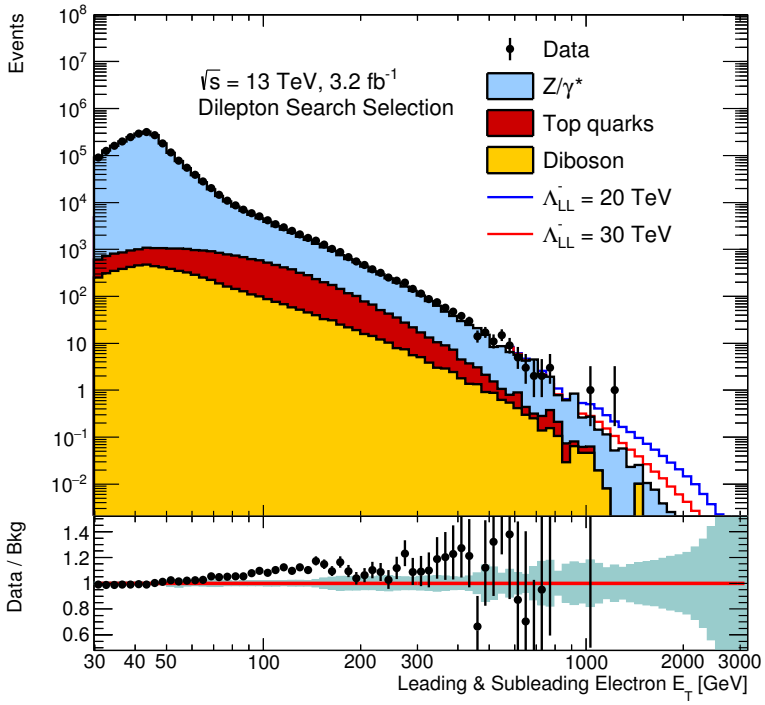


Figure 5.24: $\sqrt{s} = 13$ TeV E_T distribution for both leading and subleading electrons.

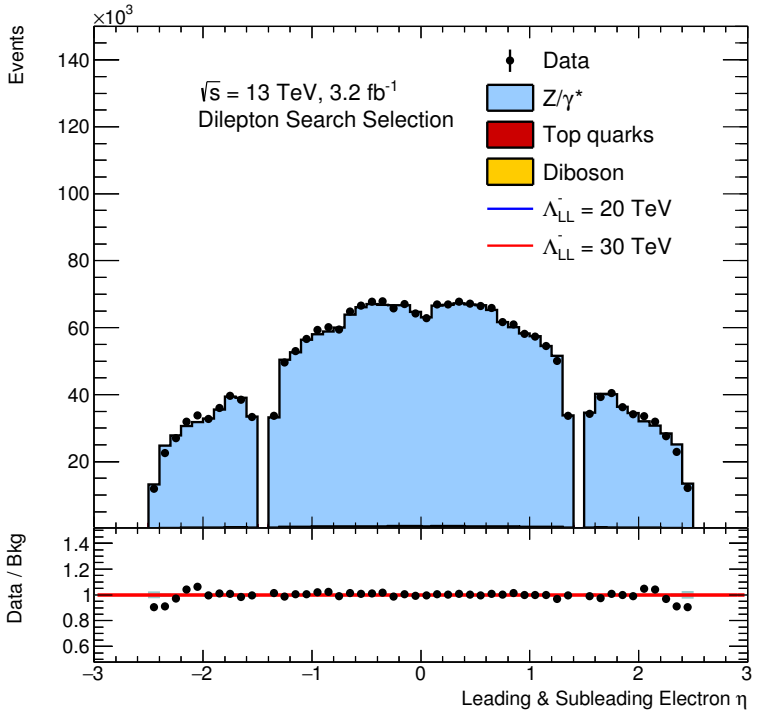


Figure 5.25: $\sqrt{s} = 13$ TeV η distribution for both leading and subleading electrons.

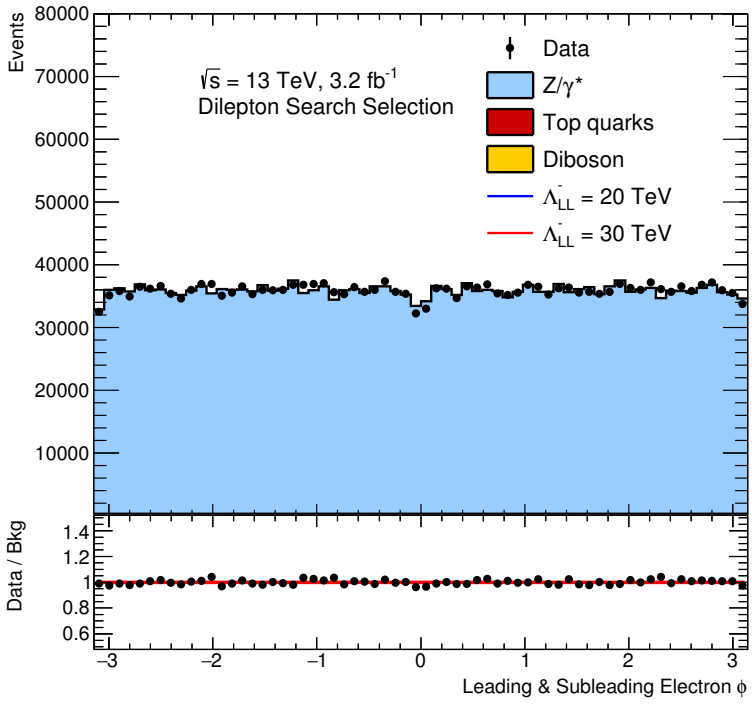


Figure 5.26: $\sqrt{s} = 13$ TeV ϕ distribution for both leading and subleading electrons.

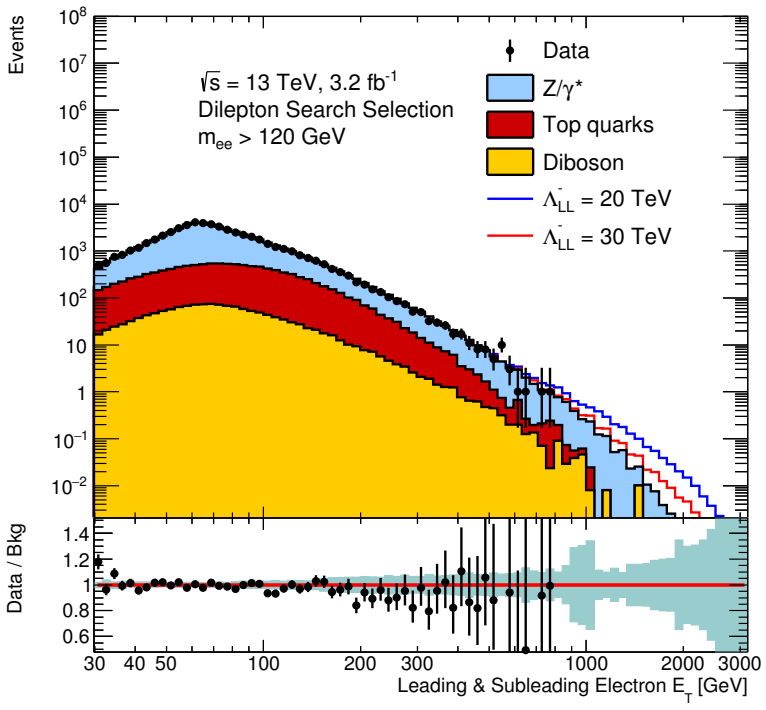


Figure 5.27: $\sqrt{s} = 13$ TeV leading and subleading E_T distribution for $m_{ee} > 120$ GeV.

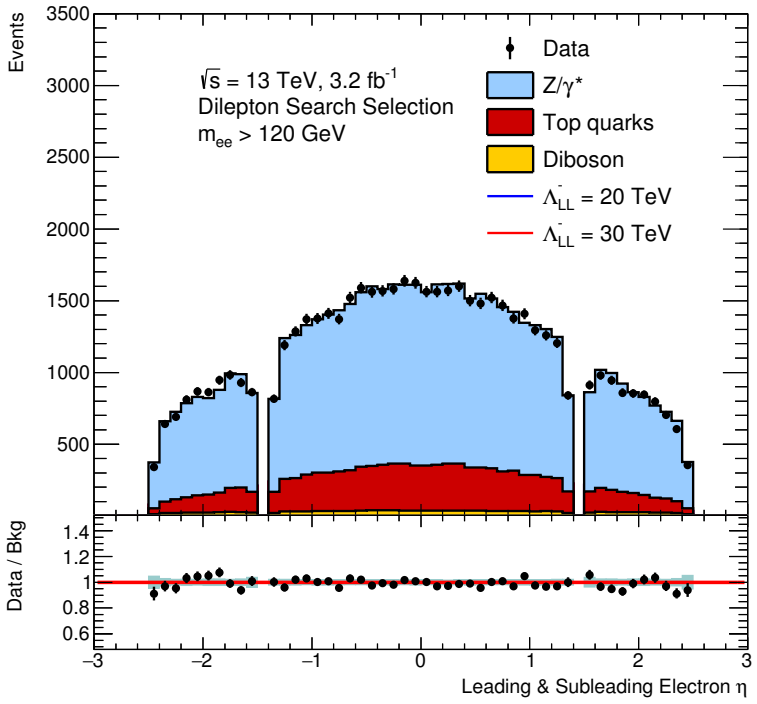


Figure 5.28: $\sqrt{s} = 13$ TeV leading and subleading η distribution for $m_{ee} > 120$ GeV.

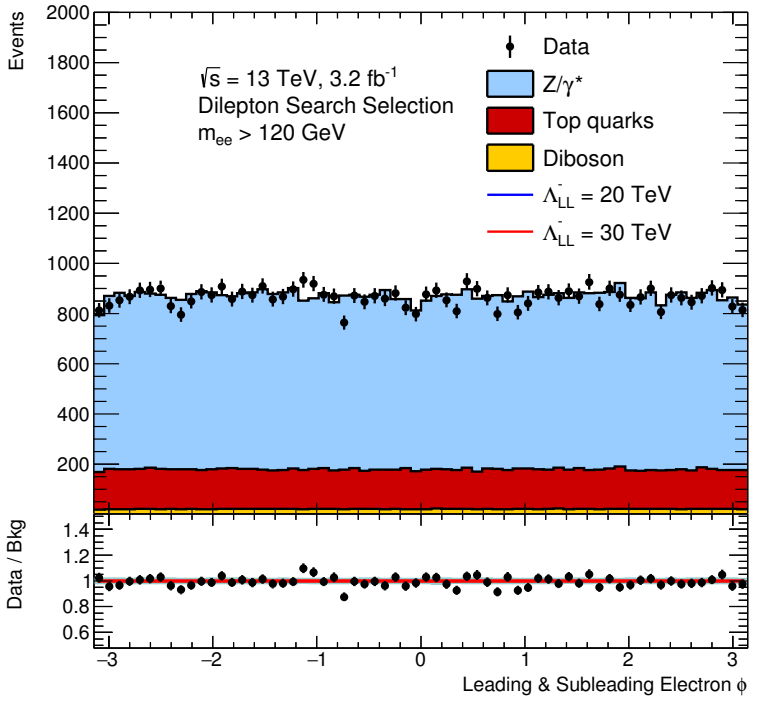


Figure 5.29: $\sqrt{s} = 13$ TeV leading and subleading ϕ distribution for $m_{ee} > 120$ GeV.

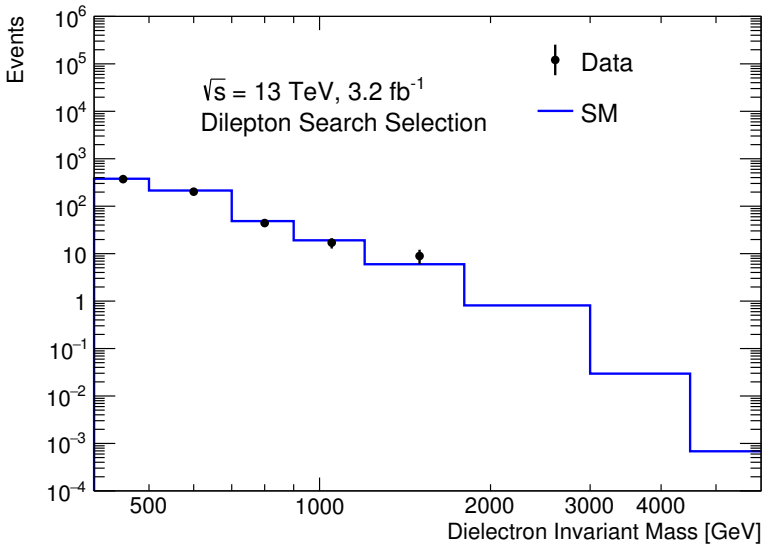


Figure 5.30: Mass binning used for the CI search and limit setting phase (400 – 500, 500 – 700, 700 – 900, 900 – 1200, 1200 – 1800, 1800 – 3000, 3000 – 4500, and 4500 – 6000 GeV). The sum of all SM backgrounds is shown after the full dielectron selection alongside the observed data.

"There are things known and there are things unknown,
and in between are the doors of perception."

–Aldous Huxley

6

Systematic Uncertainties

In a counting experiment there are many sources that could cause a variation in the measurement. Typically these arise from uncertainties in theoretical assumptions or mis-modelling of experimental measurements. Normalising the total background estimate to data in the Z-peak invariant mass region of 80–120 GeV protects each analysis against mass independent systematic uncertainties as any overall constant scale factors cancel out. This normalisation procedure is found to agree with the expected integrated luminosity within 2% for both analyses. Mass dependent systematic uncertainties still need to be considered as the shape of the discriminating variable is affected. This chapter reviews the considered systematic uncertainties on the invariant mass distribution. Firstly a discussion of the theoretical then experimental systematics is given with any differences in the uncertainties used in each of the analyses highlighted. Each uncertainty is investigated as a function of true dielectron invariant mass for signal and background where applicable. The ratios of the nominal and shifted reconstructed invariant mass distributions are used to quantify the uncertainties used in the statistical analysis and are presented for the separate ADD and CI searches.

6.1 Theoretical Uncertainties

Uncertainties on theoretical assumptions made in the analysis can have a large effect on the invariant mass shape. These uncertainties are applied to the DY MC cross-section to quantify the effect on the invariant mass shape. Theoretical uncertainties are not applied to signal, however, the effect on the signal acceptance $\times$ efficiency is investigated where relevant.

Z/γ^* Cross-Section

As the DY MC is normalised within the Z-region to data, the uncertainty on the Z/γ^* cross-section in this region is not applied. However, as it is not known how the signal shape should change by this normalisation, a uniform uncertainty of 4% due to the uncertainty on the Z/γ^* cross-section is applied to each signal sample. This uncertainty was obtained using the VRAP program [77] at NNLO by varying the PDF, scale and α_s [4, 78].

PDF Variation

The largest systematic uncertainty at high invariant mass arises from the variation in the PDFs and the effect on the DY cross-section. Each nominal PDF (as described in 4.2.1) has a set of independent parameters associated with it (eigenvectors), which can be varied in orthogonal directions to quantify the systematic uncertainties associated with variations in PDF. This was done for 20 eigenvectors, returning 20 up and 20 down cross-section variations vs truth mass. In the ADD (CI) searches these were grouped into 4 (7) similar mass dependent bundles to optimise computational time in the statistical framework.

As the ADD search is performed in a single high mass-bin (described in 7.4.1), the variation in the acceptance $\times$ efficiency of the search region is studied. An average systematic uncertainty of 6% was found in the acceptance $\times$ efficiency for the ADD MC samples, and is included as a nuisance parameter in the statistical analysis. This study is conducted using the largest uncertainty (discussed in 1.1.3), the PDF variation theory systematic, making the results most conservative.

PDF Choice

The uncertainty due to the choice of PDF was investigated by assessing the variations of other PDFs compared to the nominal. In the 2012 analysis the CT10NNLO [58], NNPDF2.3 [72], ABM11 [79] and HERAPDF1.5 [80] PDF sets are compared with the nominal PDF variation uncertainty. Only ABM11 is found to vary outside the nominal PDF variation and as such the difference is included as a further uncertainty. A similar process is conducted in the 2015 analysis with variations in the PDF sets recommended by the PDF4LHC forum [81], namely MMHT14 [82], NNPDF3.0 [72], ABM12 [79] and JR14 [83]. NNPDF3.0 is found to have the largest variation outside the nominal PDF variation and the difference is included as a further uncertainty.

PDF Scale

In order to quantify the uncertainties due to the variation of the PDF scale and QCD coupling strength, α_S, the DY cross-section is calculated as a function of truth mass by simultaneously varying the renormalisation and factorisation scales up and down by a factor of two and using α_S values of 0.118 ± 0.003. In run 2 the PDF4LHC [81] recommended a variation of $\pm$ 0.0015, but due to the high mass reach of this search it was deemed appropriate to stick with the nominal conservative uncertainty of $\pm$ 0.003. A variation $<$ 3% throughout the mass spectrum was observed due to the scale variation, as such the conservative estimate had negligible impact on results.

Electroweak Higher-order Corrections

The higher order EW corrections are re-calculated using an alternative programme, SANC [84]. The systematic uncertainty on the EW corrections corresponds to the difference in these calculations from the nominal.

Photon-Induced

The PI background was estimated in the $\sqrt{s}$ = 8 TeV analysis through MC, and as a PI k-factor in the $\sqrt{s}$ = 13 TeV analysis, as described in Section 4.2.2. To assign an uncertainty in the $\sqrt{s}$ = 8 TeV ADD search, a conservative 100% error to the PI contribution is investigated. The systematic uncertainty is calculated as the ratio between

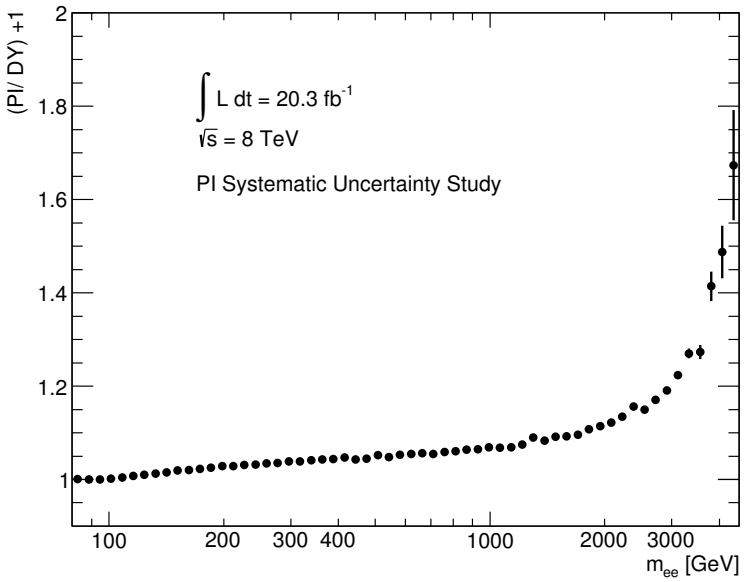


Figure 6.1: Systematic uncertainty of including PI process.

PI and DY MC to quantify the effect of not including this MC background and is shown in Figure 6.1. This could also then be performed for different angular regions and as a function of invariant mass for the parallel search, where the uncertainty on the k-factor had no angular dependence.

In the 2015 analysis, the central value of the k-factor is used as the PI estimate with the upper and lower estimate from the cross-section calculation as a systematic uncertainty. While the PI systematic is larger in the $\sqrt{s} = 8$ TeV analysis compared to the $\sqrt{s} = 13$ TeV, the MC generation was required due to the need of angular information in the $\sqrt{s} = 8$ TeV CI search.

6.2 Experimental Uncertainties

Experimental systematics are applied to both signal and background, with the exception of the multi-jet and W+jets systematic which is applied only to the relevant background.

Efficiency

Systematic uncertainties due to the trigger, isolation, identification and reconstruction efficiencies are provided by the ATLAS electron/photon performance group. For each derived SF used to shift MC efficiencies to that seen in data (see section 4.1.2) an up and down variation by 1σ is used to assess the uncertainty of the SF. An additional 2% uncertainty is assigned to the electron identification SF at high E_T to accommodate changes in the shapes of variables used in the likelihood [56].

Energy Scale/Resolution

Systematic uncertainties due to energy scale and resolution smearing are provided by the ATLAS electron/photon performance group. For each derived smearing (see section 4.1.2) an up and down variation is used to assess the uncertainty [55].

Multi-jet/W+jets

The data-driven method used to estimate the fake electron contributions to the event selection (as explained in Section 4.2.5) uses two (three) methods to ascertain an uncertainty in the $\sqrt{s} = 8$ TeV ($\sqrt{s} = 13$ TeV) analysis. Firstly, the envelope of the ensemble of fits used in the extrapolation. Secondly, by setting the real efficiency = 1 and performing the method again the ratio to the nominal could then be used as a mass dependent uncertainty. At $\sqrt{s} = 13$ TeV an additional uncertainty is studied using the ratio (R) of (opposite sign − same sign) / (opposite sign + same sign) events. If the event came from the dominant DY or $t\bar{t}$ contributions, the electrons would have opposite charges and R would tend towards 1. If the event contained a fake, there would be no preference for the fake to be of positive or negative electric charge and R would tend to 0. This value was determined using the integral of events above 120 GeV and was found to be 11%. The combination of uncertainties were then summed in quadrature to give an overall value.

The $\sqrt{s} = 8$ TeV analysis found a maximal variation of $\pm 20\%$ in the estimate, as such a flat maximal 20% variation was used as an uncertainty. Due to the reduced contribution of fake events in the $\sqrt{s} = 13$ TeV analysis with the introduction of the likelihood identification, the variations had a much larger effect (100% at 1 TeV). Despite this uncertainty being extremely large, it is only on the smallest of the backgrounds and makes negligible difference to the total uncertainty.

Beam Energy

The DY cross-section is calculated using VRAP with varying beam energies (up/down by 0.65% [85]). The ratio between the varied and nominal invariant mass distributions are used to quantify the uncertainty in yields. This calculation was not performed in the CI search. Adding in quadrature the $\sqrt{s} = 8$ TeV beam energy uncertainty at a dielectron invariant mass of 3 TeV (5%) to the $\sqrt{s} = 13$ TeV total uncertainty at 3 TeV, an increase of 0.7 (1.3) % is observed in the background (signal) uncertainty.

Charge Mis-Identification

The 2012 analysis included an opposite sign requirement. The mis-identification rate was assessed and used as an uncertainty in the analysis. An electrons electric charge can be mis-identified in two ways. Firstly if the electron is very energetic, its track will have minimal curvature, thus difficult to determine the electrons charge. Secondly, if the electron emits a hard photon through Bremsstrahlung, which in turn pair produces, one of the pair produced electrons could be selected by mistake if it has sufficient E_T. The electron charge mis-identification systematic was investigated by adding 100% additional charge mis-identification at truth level throughout the DY sample. Comparing this to the nominal selection an uncertainty was calculated to have a 3% effect at high invariant mass.

Monte Carlo Statistics

The statistical error on all MC was also taken into account for both background and signal. By comparing the nominal to the shifted mass distribution the systematic uncertainty is assessed. This is largest in the ADD signal due to the smaller number of mass bins used in each process simulation.

6.3 ADD Analysis

Each relevant systematic uncertainty is applied to the ADD signal and SM background reconstructed invariant mass spectrums. The ratio of shifted and nominal distributions are presented in Table 6.1 for 3 benchmark dielectron masses of 1 TeV (2 TeV) [3 TeV].

Table 6.1: Summary of the systematic uncertainties in the expected number of events at a dielectron invariant mass of 1 TeV (2 TeV) [3 TeV] in the $\sqrt{s} = 8$ TeV ADD search.

Source	Dielectrons	
	Signal	Background
Normalization	4.0% (4.0%) [4.0%]	NA
PDF Variation	6.0% (6.0%) [6.0%]	5.0% (11.0%) [30.0%]
PDF Choice	NA	1.0% (7.0%) [22.0%]
PDF Scale	NA	1.0% (3.0%) [5.0%]
EW Corrections	NA	1.0% (2.0%) [4.0%]
Photon-Induced	NA	7.0% (12.0%) [21.0%]
Efficiency	1.0% (2.0%) [3.0%]	1.0% (2.0%) [3.0%]
Scale/Resolution	1.2% (2.4%) [5.0%]	1.2% (2.4%) [5.0%]
Multi-jet/W+jets	NA	<1.0% (<1.0%) [<1.0%]
Beam Energy	1.0% (3.0%) [5.0%]	1.0% (3.0%) [5.0%]
Charge MisID	1.2% (2.0%) [2.9%]	1.2% (2.0%) [2.9%]
Statistical	3.0% (3.0%) [3.0%]	<1.0% (<1.0%) [<1.0%]
Total	8.1% (9.2%) [11.3%]	9.1% (18.7%) [44.0%]

6.4 Contact Interaction Analysis

Each relevant systematic uncertainty is applied to the CI signal and SM background reconstructed invariant mass spectrums. The ratio of shifted and nominal distributions are presented in Table 6.2 for 3 benchmark dielectron masses of 1 TeV (2 TeV) [3 TeV].

Table 6.2: Summary of the systematic uncertainties in the expected number of events at a dielectron invariant mass of 1 TeV (2 TeV) [3 TeV] in the $\sqrt{s} = 13$ TeV CI search.

Source	Dielectrons	
	Signal	Background
Normalisation	4.0% (4.0%) [4.0%]	NA
PDF Variation	NA	5.1% (8.6%) [13.5%]
PDF Choice	NA	<1.0% (<1.0%) [<1.0%]
PDF Scale	NA	1.2% (1.8%) [2.3%]
EW Corrections	NA	0.8% (2.1%) [3.9%]
Photon-Induced	NA	2.1% (3.3%) [5.4%]
Efficiency	5.4% (5.4%) [5.4%]	5.4% (5.4%) [5.4%]
Scale/Resolution	2.9% (3.7%) [5.4%]	2.9% (3.7%) [5.4%]
Multi-jet/W+jets	NA	<1.0% (<1.0%) [<1.0%]
Statistical	<1.0% (<1.0%) [<1.0%]	<1.0% (<1.0%) [<1.0%]
Total	7.3% (7.7%) [8.6%]	8.6% (11.7%) [17.0%]

"Data do not give up their secrets easily. They must be tortured to confess."

–Jeff Hopper, Bell Labs

7

Statistical Analysis

As it is only possible to infer the production mechanism from measuring the final state particles of a selected event, a method is required to quantify the degree of belief that new physics is present. This chapter describes the statistical analysis performed on the selected events. Firstly a brief overview of the Bayesian method used to quantify discovery and set lower exclusion limits of a model is given. Followed by a discussion of the choice of priors used. Finally the counting experiments used in each search including the optimisation of the ADD search region is explained.

7.1 Bayesian Analysis

In a given invariant mass region, the number of expected events, μ, is calculated by the sum of SM background events, n_b, and the expected contribution from new physics signal, n_s;

$$\mu = n_s(\Theta, \bar{\Omega}) + n_b(\bar{\Omega})), \tag{7.1}$$

where $\bar{\Omega}$ represents the set of Gaussian nuisance parameters corresponding to the systematic shifts in number of expected signal and background events. The variable Θ corresponds to the parameter of interest; in the ADD search this is the string scale, M_{S}, and in the CI search this is the compositeness energy scale, Λ. The likelihood of observing $\bar{n}$ events

given the new physics parameter Θ and nuisance parameters $\bar{\Omega}$ is found using the product of Poisson probabilities across N bins;

$$\mathcal{L}(\bar{n}|\Theta,\bar{\Omega}) = \prod_{k=1}^{N} \frac{\mu_k^{n_k} e^{-\mu_k}}{n_k!}, \tag{7.2}$$

where n_k is the number of observed events in data in the k^{th} bin. Employing Bayes' theorem, the posterior probability density for the parameter Θ given $\bar{n}$ observed events is

$$P(\Theta|\bar{n}) = \frac{1}{Z}\mathcal{L}_{\mathcal{M}}(\bar{n}|\Theta)P(\Theta), \tag{7.3}$$

where $\mathcal{L}$ is the marginalised likelihood, which is the likelihood after all nuisance parameters have been integrated out, leaving a distribution dependent on the parameter of interest only (Θ). Z corresponds to a normalisation constant. P(Θ) is the prior probability, chosen to be uniform and positive.

After event selection has been completed and an absence of signal found, a limit on the parameter of interest to a 95% credibility level (C.L.) can be determined. The 95% C.L. limit is found by satisfying;

$$\int_0^{\Theta_{lim}} P(\Theta|\bar{n})d\Theta = 0.95, \tag{7.4}$$

for Θ_{Lim}. Using the statistical software framework; Bayesian Analysis Toolkit (BAT) [86], these calculations are made with a Markov Chain Monte Carlo technique to integrate out the dependence on the nuisance parameters.

7.1.1 Choice of Prior

As the relative sizes of the pure and interference terms are not known *a priori*, the analyses interpret the result twice using two separate priors for each new physics model. A positive prior flat in either $1/M_S^4$ or $1/M_S^8$, motivated by Equation 1.4, is used for the ADD search. While the ADD analysis investigates both priors for consistency and comparison to previous results, the interference effects between the DY and virtual graviton processes are small due to the dominance of the gluon-induced over quark-induced production mechanisms. This is shown in Figure 7.1, where the turn on for each samples deviation from

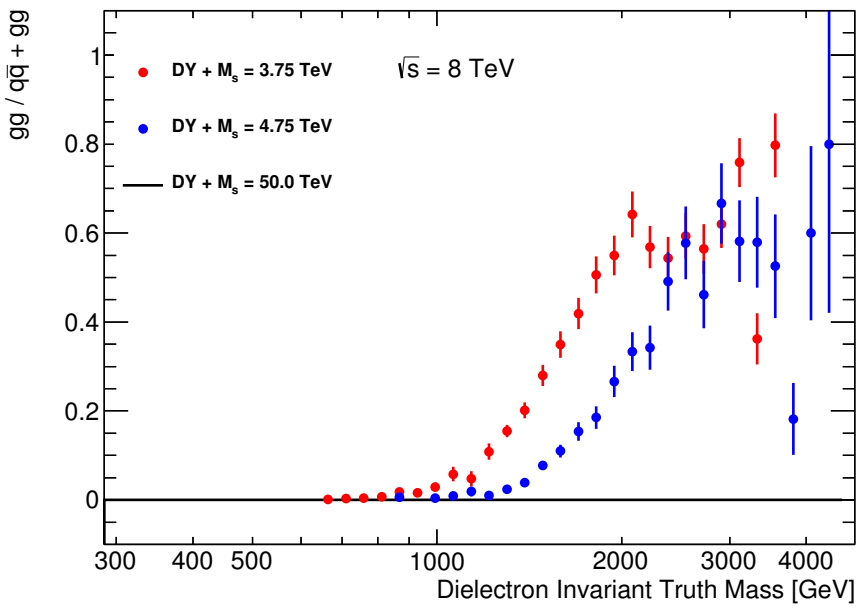


Figure 7.1: Fraction of ADD graviton production through gg and $q\bar{q}$ initiated processes.

DY becomes gluon-gluon dominated, making the $1/M_S^8$ interpretations more appropriate. The $M_S = 50$ TeV sample, being essentially DY-only in the kinematic region being probed is entirely $q\bar{q}$ initiated and is flat at $gg/q\bar{q} + gg = 0$.

Similarly, a flat positive prior in either $1/\Lambda^2$ or $1/\Lambda^4$, motivated by equation 1.7 is used for the CI search. As the interference term begins to dominate at higher Λ signal values the $1/\Lambda^2$ prior is deemed most appropriate.

7.2 Discovery Statistics

In order to check the consistency between the data and the SM expectation, the log-likelihood ratio (LLR) between the signal+background and pure background hypotheses obtained in the data is compared to the results from 1000 SM only pseudo-experiments. According to the Neyman–Pearson Lemma this is the best discriminant between signal+background and background only hypotheses. The LLR is calculated as;

$$LLR = -2\frac{\mathcal{L}(\bar{n}|\Theta,\bar{\Omega})}{\mathcal{L}(\bar{n}|0,\bar{\Omega})}. \tag{7.5}$$

To construct the background-only likelihood for data, Eq. 7.2 is used by taking μ_k

as the number of SM only events from simulation and n_k as the number of observed events, in each mass bin k. The same calculation is done for each of the SM only pseudo-experiments by varying the background expectation for values of n_k according to a Poisson distribution. The signal+background likelihood is constructed by taking μ_k as the number of signal+background events and evaluated at the M_S or Λ value that maximises the likelihood for the respective search.

The p-value corresponds to the probability of observing a fluctuation at least as signal-like in data, given that only background exists. The convention in particle physics is that a p-value $< 1.35 \times 10^{-3}$ corresponds to evidence for a signal, whereas a p-value $< 2.87 \times 10^{-7}$ is considered a discovery. These values are the one-sided integrals of the tails of a unit Gaussian distribution beyond 3σ and 5σ, respectively.

A p-value is derived by taking the ratio of the number of pseudo-experiments (N_{PE}) with a LLR more signal-like than observed in data relative to the total number of the pseudo-experiments (N_{PE}^{Tot}):

$$p = \frac{N_{PE}(LLR < LLR_{data})}{N_{PE}^{Tot}}. \tag{7.6}$$

7.3 Signal Parameterisation

In the posterior calculation, the parameter of interest that maximises the likelihood might be found to have any value. It is only possible to generate a limited number of MC samples, which can be used to relate the number of expected events (N^{exp}) to a given M_S or Λ. To provide more values of N^{exp}, a fit of the number of expected events in each bin is performed using the MC samples according to

$$N^{exp}(M_S) = a_0 + \frac{a_1}{M_S^4} + \frac{a_2}{M_S^8}, \tag{7.7}$$

for the ADD search, and

$$N^{exp}(\Lambda) = b_0 + \frac{b_1}{\Lambda^2} + \frac{b_2}{\Lambda^4}. \tag{7.8}$$

for the CI search. The a_0 parameter is obtained using an ADD signal sample with a very high M_S value, 50 TeV, where only a DY contribution is present in the search bin. The

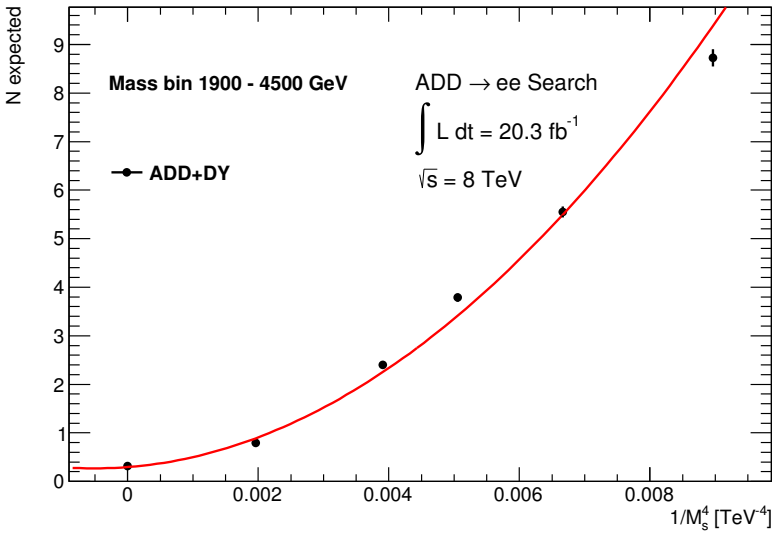


Figure 7.2: Number of expected events in the electron channel in the ADD search region as a function of $1/M_S^4$.

b_0 parameter is obtained using a DY only sample, estimated using the same generator as the CI signal. These DY values are required to produce the parameterisations due to the combined production of signal and DY background in the MC samples. Once the parameterisation is created the 0th parameter (SM DY value) can be set to zero, leaving a signal only function used in the new physics estimate. Figure 7.2 shows an example of the SHERPA DY + ADD signal parameterisation with an invariant mass cut of 1900 GeV.

Weighting the signal MC using the systematic uncertainties and parameterising in the same manner as the nominal number of expected events, signal parameterisation for each systematic can be provided to be used in the statistical interpretation. A full display of these signal uncertainties for the ADD search region at 1900 GeV can be seen in Figure 7.3 (ZXS and BE correspond to the normalisation and beam energy uncertainties, respectively).

Figures 7.4, 7.5, 7.6 and 7.7 show the CI signal parameterisation in each of the search invariant mass regions for the LL constructive case, the destructive case can be found in Appendix C. Similar parameterisations are created for the other models under investigation. In the same manner as the ADD search, each CI signal parameterisation has a corresponding set of systematics parameterisations which can be found in Appendix C.

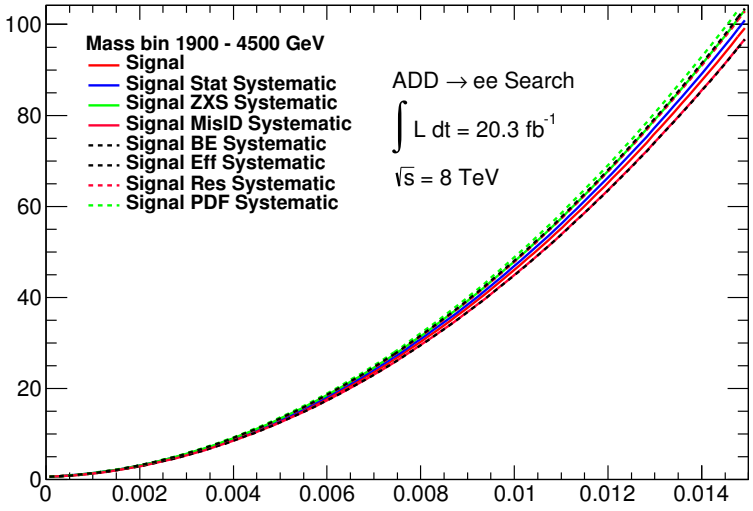


Figure 7.3: Systematically varied number of expected events in the electron channel for the ADD search region as a function of $1/M_S^4$.

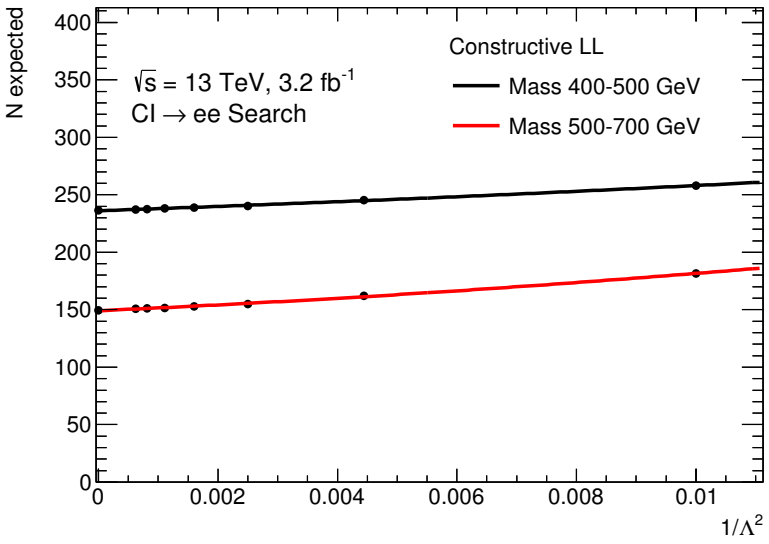


Figure 7.4: Number of expected events in the electron channel for the CI LL constructive interference 400–500 and 500–700 GeV search regions as a function of $1/\Lambda^2$.

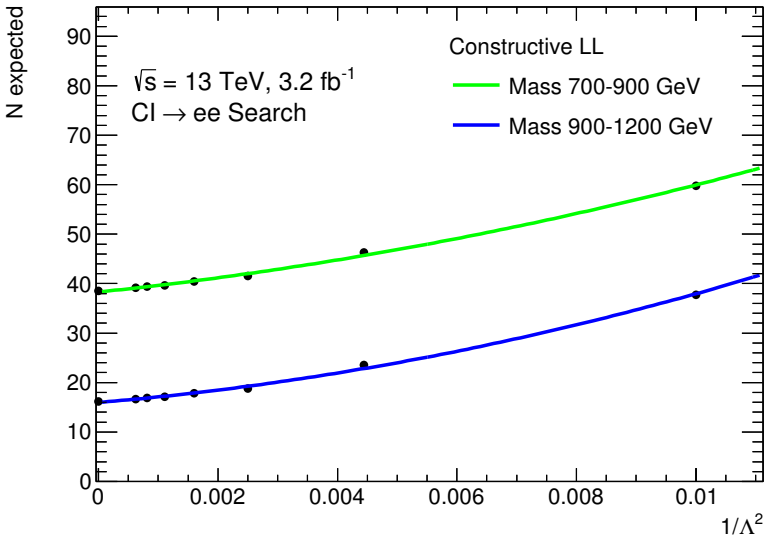


Figure 7.5: Number of expected events in the electron channel for the CI LL constructive interference 700–900 and 900–1200 GeV search regions as a function of $1/\Lambda^2$.

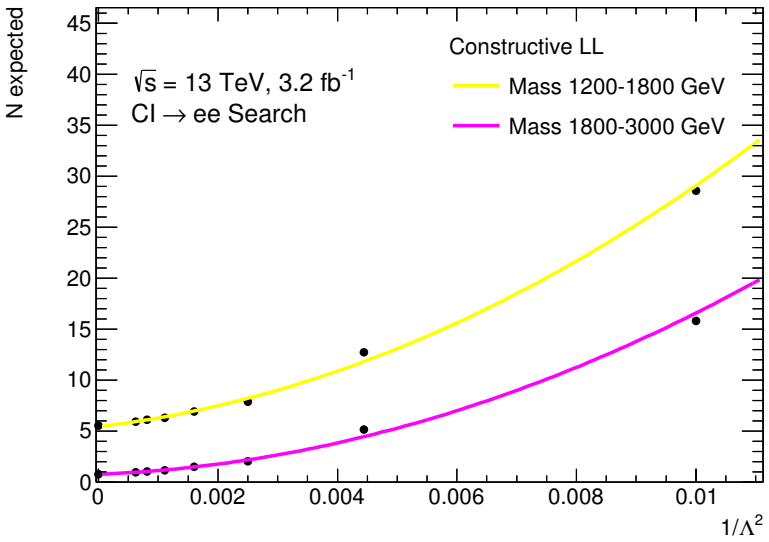


Figure 7.6: Number of expected events in the electron channel for the CI LL constructive interference 1200–1800 and 1800–3000 GeV search regions as a function of $1/\Lambda^2$.

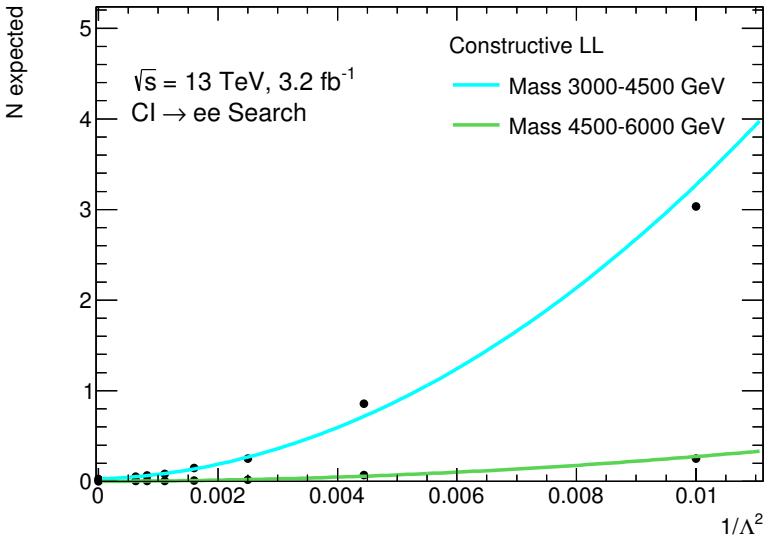


Figure 7.7: Number of expected events in the electron channel for the CI LL constructive interference 3000–4500 and 4500–6000 GeV search regions as a function of $1/\Lambda^2$.

7.4 Counting Experiment

7.4.1 ADD Search

The ADD search uses a single high invariant mass bin counting experiment to search for new physics. Since the ultraviolet cutoff imposed at a given M_S is caused by limited theory beyond this scale, a single bin approach is preferred which tests the number of expected signal events rather than the shape of the signal. Due to the sharp turn on of signal at high mass the lower mass threshold of the single bin is optimised such that the expected lower limit is maximised.

Optimisation

A one bin search is performed for ADD signals above an invariant mass threshold. The lower bound in invariant mass (m_{ee}^{min}) of the search region was selected by calculating the expected limit on M_S as a function of the lower mass threshold. The lower edge of the single mass bin was varied from 1000 GeV to 2200 GeV in 100 GeV steps and the expected limits were determined. This was done independently for the electron and then again for a combined lepton search (ee +$\mu\mu$), using the inputs provided by Marc Cano Bret and

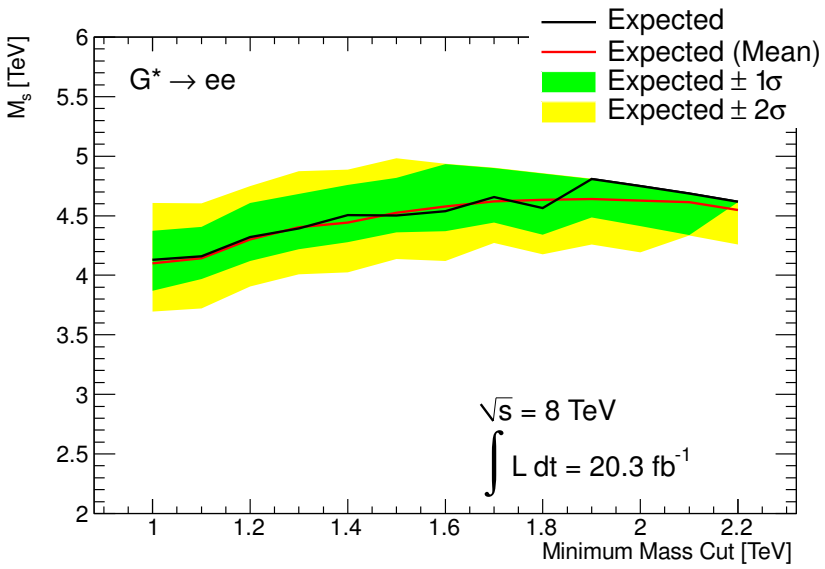


Figure 7.8: Expected median (black) and mean (red) limit on M_S in the ADD GRW model determined as a function of minimum mass cut (m_{ee}^{min}) used to define the search region using a $1/M_S^4$ for the dielectron channel.

Tülin Varol. The expected limits were computed by generating a set of 1000 SM-like PEs, in which no ADD signal was present. Then the full Bayesian method calculation was carried out for each of the PEs to extract 1000 separate 95% C.L. limits. To determine the optimal value of $m_{\ell\ell}^{min}$, two definitions of the limit on M_S are used: the mean and the median of the PE expected limit distribution. At the highest values of $m_{\ell\ell}^{min}$ the expected limit PE distribution collapses to a few discrete results rather than forming a continuous distribution. Using the median definition, the limit can change value quickly as the median jumps from one value to another. To avoid this behaviour the limit is taken to be the mean of the expect limit PE distribution. These are shown for the dielectron and dilepton ADD decay channels (denoted as G*) under both priors in Figures 7.8, 7.9, 7.10 and 7.11. The mean limit definition is seen to be smooth in comparison to the median limit as a function of $m_{\ell\ell}^{min}$. Considering the expected limits, using both the $1/M_S^4$ and $1/M_S^8$ priors, the lower cut was chosen to be 1900 GeV for all channels (a similar study was conducted in the muon channel) and used as the search region in the ADD analysis.

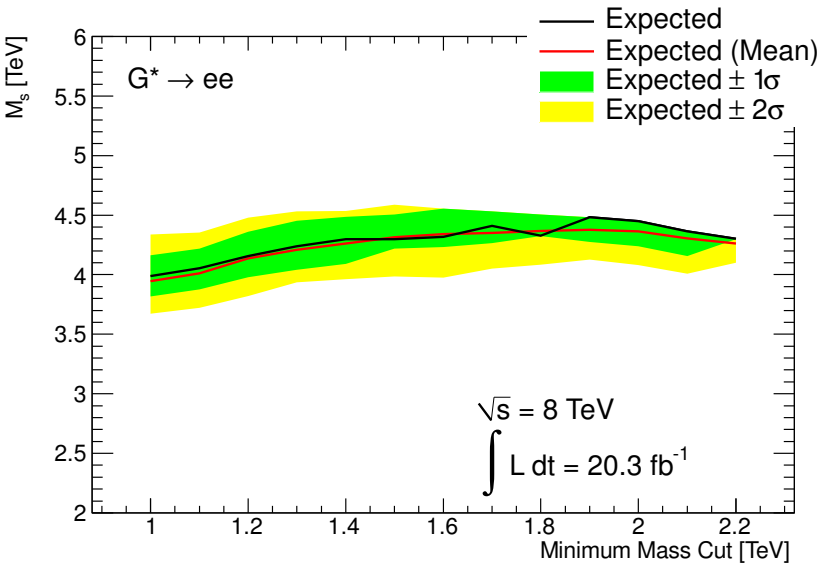


Figure 7.9: Expected median (black) and mean (red) limit on M_S in the ADD GRW model determined as a function of minimum mass cut (m_{ee}^{min}) used to define the search region using a $1/M_S^8$ for the dielectron channel.

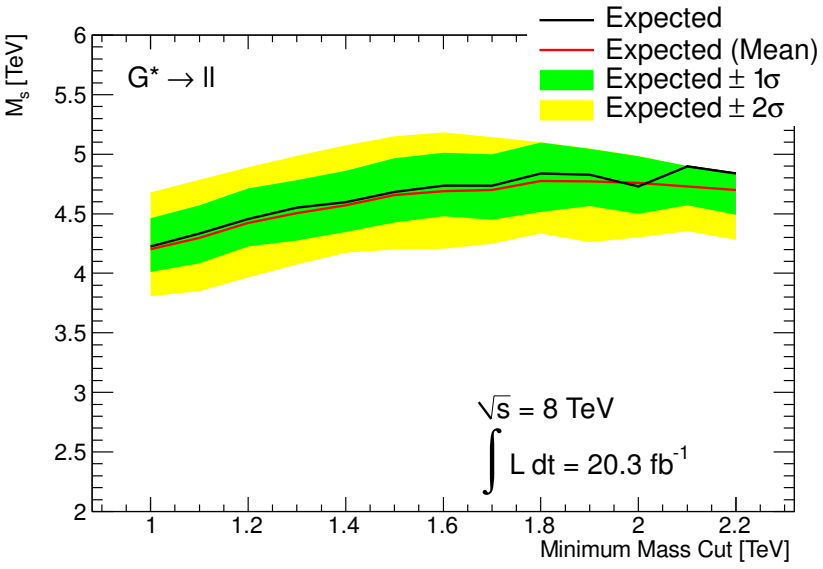


Figure 7.10: Expected median (black) and mean (red) limit on M_S in the ADD GRW model determined as a function of minimum mass cut ($m_{\ell\ell}^{min}$) used to define the search region using a $1/M_S^4$ for the dilepton channel.

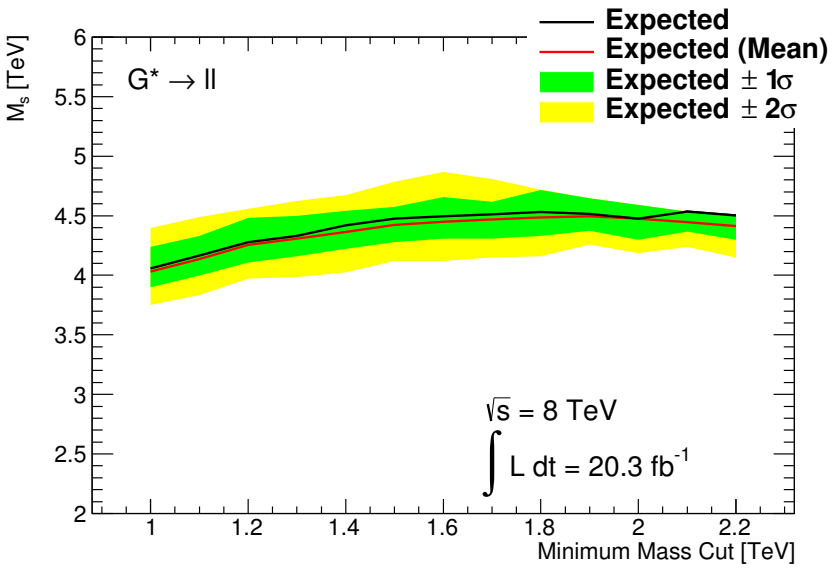


Figure 7.11: Expected median (black) and mean (red) limit on M_S in the ADD GRW model determined as a function of minimum mass cut ($m_{\ell\ell}^{min}$) used to define the search region using a $1/M_S^8$ for the dilepton channel.

7.4.2 Contact Interaction Search

Due to the broad nature of the expected CI signal, the CI search uses eight large invariant mass bins in which to perform a counting experiment. These regions have bin edges of 400, 500, 700, 900, 1200, 1800, 3000, 4500 and 6000 GeV.

"The absence of evidence is not the evidence of absence."

–Carl Sagan

8 Results

After selecting potential signal candidates and determining the search regions, the statistical analysis can be performed to quantify any excess. In the absence of an excess, limits can be set on the parameter of interest. This chapter focuses on the results of the ADD and CI searches.

8.1 ADD Analysis

8.1.1 Search

The LLR distributions for data and the 1000 PEs are shown in Figures 8.1, 8.2 for the ADD GRW formalism in electron channel under the $1/M_S^4$ and $1/M_S^8$ priors. Due to the single search bin and the small number of expected events, when varying the number of simulated observed events in a PE according to Poisson distribution, only a few values of M_S are found to maximise the likelihood. This is the cause of the highly discretised nature of the ADD LLR and limit distributions. The derived p-values in Table 8.1 show that there is no significant evidence for LED in data with a dielectron or dilepton final state. These dilepton results relied upon the inputs Marc Cano Bret and Tülin Varol.

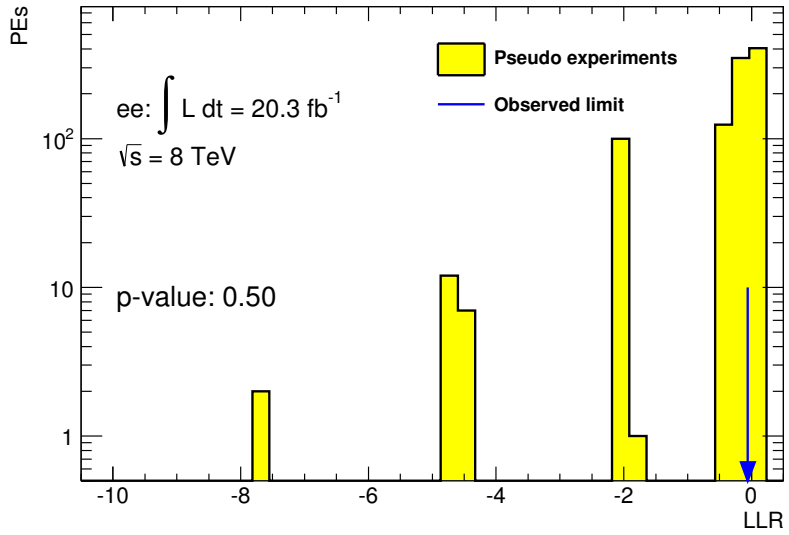


Figure 8.1: Negative log-likelihood distribution resulting from pseudo-experiments with fluctuations on the number of expected SM-only events in the dielectron channel for the GRW ADD model with uniform positive prior in $1/M_S^4$.

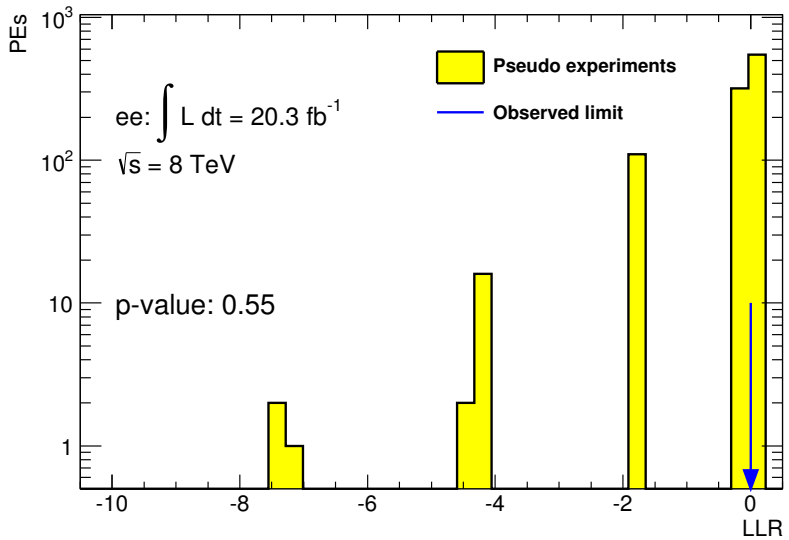


Figure 8.2: Negative log-likelihood distribution resulting from pseudo-experiments with fluctuations on the number of expected SM-only events in the dielectron channel for the GRW ADD model with uniform positive prior in $1/M_S^8$.

Table 8.1: Derived p-values for the GRW ADD dielectron and dilepton searches using uniform positive priors in $1/M_S^4$ and $1/M_S^8$.

p-value [%]	$1/M_S^4$	$1/M_S^8$
ee	50	55
$\ell\ell$	19	53

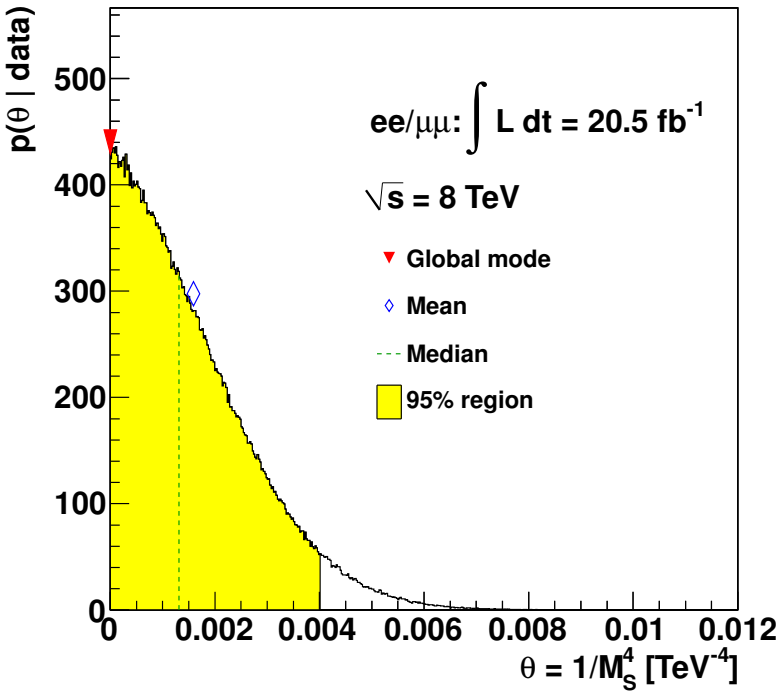


Figure 8.3: Posterior probability density distributions for the GRW ADD model in the dielectron channel, with a uniform positive prior in $1/M_S^4$.

8.1.2 Limit Setting

In an absence of signal, lower exclusion limits are set on the scale M_S. The posterior probability density distributions for data using the $1/M_S^4$ and $1/M_S^8$ priors in the electron channel are shown in Figures 8.3 and 8.4. The Θ_{Lim} value satisfying Equation 7.4 for these posterior distributions is the observed 95% C.L. on the ADD GRW model. The same procedure is conducted on 1000 PEs resulting in 1000 expected limits, the median of which is taken as the expected limit on the model. These limits are displayed in Figure 8.5 and Figure 8.6 for the electron channel under both prior probabilities.

The results obtained using the GRW formalism have been translated into the HLZ and Hewett formalisms using Equation 1.5. For the special case of HLZ with $n = 2$ (which can not be trivially converted from GRW results due to the dependence on dilepton invariant mass), dedicated MC was generated and subjected to the same analysis and statistical

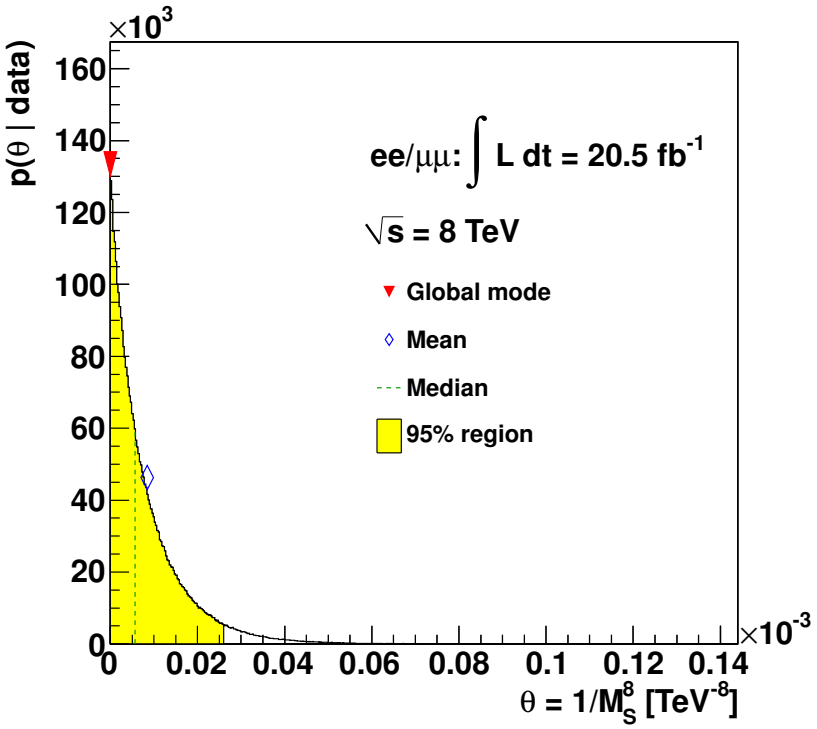


Figure 8.4: Posterior probability density distributions for the GRW ADD model in the dielectron channel, with a uniform positive prior in $1/M_S^8$.

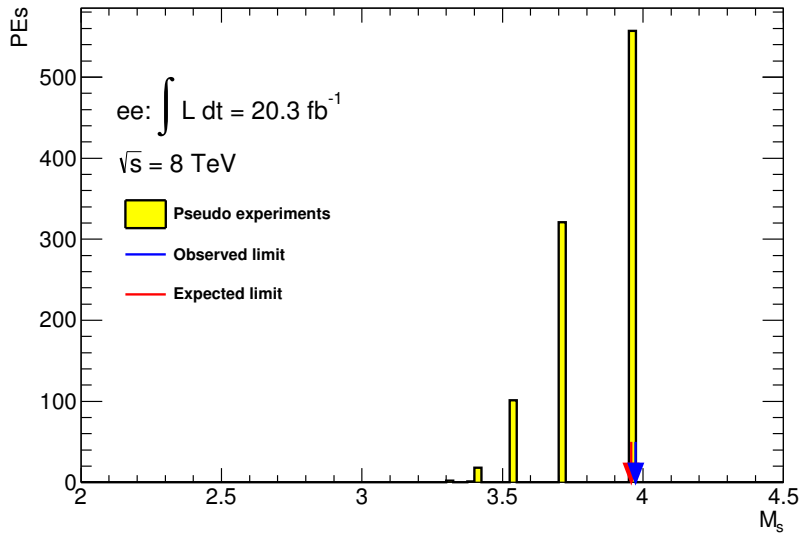


Figure 8.5: Distribution of observed and expected limit values in the dielectron channel for the GRW ADD model with uniform positive priors flat in $1/M_S^4$.

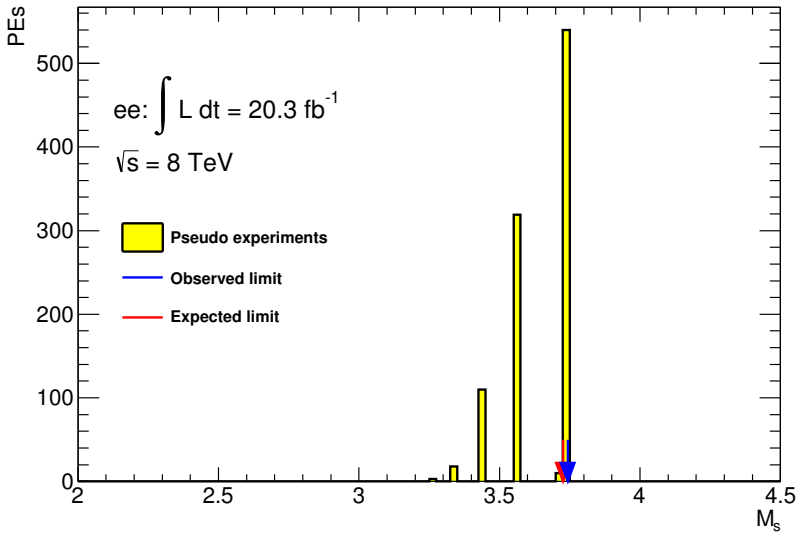


Figure 8.6: Distribution of observed and expected limit values in the dielectron channel for the GRW ADD model with uniform positive priors flat in $1/M_S^8$.

interpretation as the GRW formalism. The results range from $M_S > 3.0$–5.0 TeV, with the benchmark GRW model at $M_S > 4.0$ TeV and $M_S > 4.2$ TeV using flat prior probabilities in $1/M_S^4$ for the dielectron and dilepton final states, respectively. All results are presented in Table 8.2.

Figure 8.7 (8.8) displays the observed and expected limits for the dielectron (dilepton) final states including the 1 and 2 sigma limit bands. Included are the $\sqrt{s} = 7$ TeV results for comparison. As the HLZ $n = 2$ search has been conducted for the first time here there

Table 8.2: Observed and expected lower exclusion limits on M_S for ADD search.

Expected and observed lower limits on M_S [TeV]									
Channel	Prior	GRW	Hewett	HLZ					
				$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
Exp: ee	$1/M_S^4$	4.0	3.5	3.6	4.7	4.0	3.6	3.3	3.1
Obs: ee		4.0	3.5	3.6	4.7	4.0	3.6	3.3	3.1
Exp: ee	$1/M_S^8$	3.7	3.3	3.1	4.4	3.7	3.4	3.1	3.0
Obs: ee		3.7	3.3	3.1	4.4	3.7	3.4	3.1	3.0
Exp: $\ell\ell$	$1/M_S^4$	4.0	3.6	3.9	4.8	4.0	3.6	3.4	3.2
Obs: $\ell\ell$		4.2	3.8	4.2	5.0	4.2	3.8	3.6	3.4
Exp: $\ell\ell$	$1/M_S^8$	3.8	3.4	3.5	4.6	3.8	3.5	3.2	3.1
Obs: $\ell\ell$		4.0	3.6	3.7	4.7	4.0	3.6	3.4	3.2

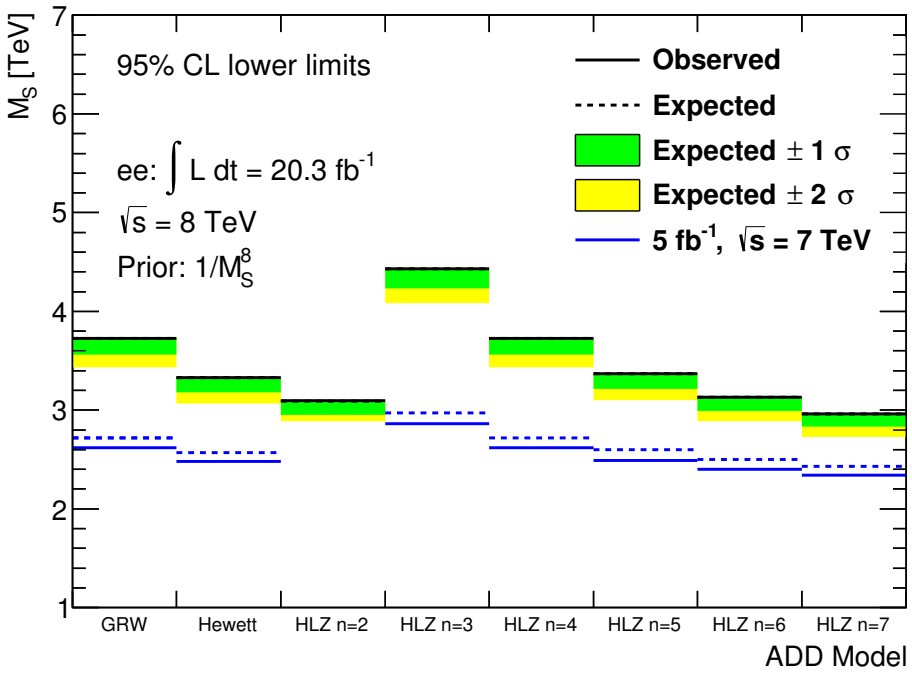


Figure 8.7: Observed and expected limit values on the ADD model searches in the dielectron channel using a uniform positive prior in $1/M_S^8$.

are no $\sqrt{s} = 7$ TeV results to compare to.

8.2 Contact Interaction Analysis

8.2.1 Search

The LLR distributions for data and the 1000 PEs are shown in Figures 8.9 and 8.10 for the CI LL constructive model in electron channel under the $1/\Lambda^2$ and $1/\Lambda^4$ priors. The derived p-values in Table 8.3 show that there is no significant evidence of signal in data.

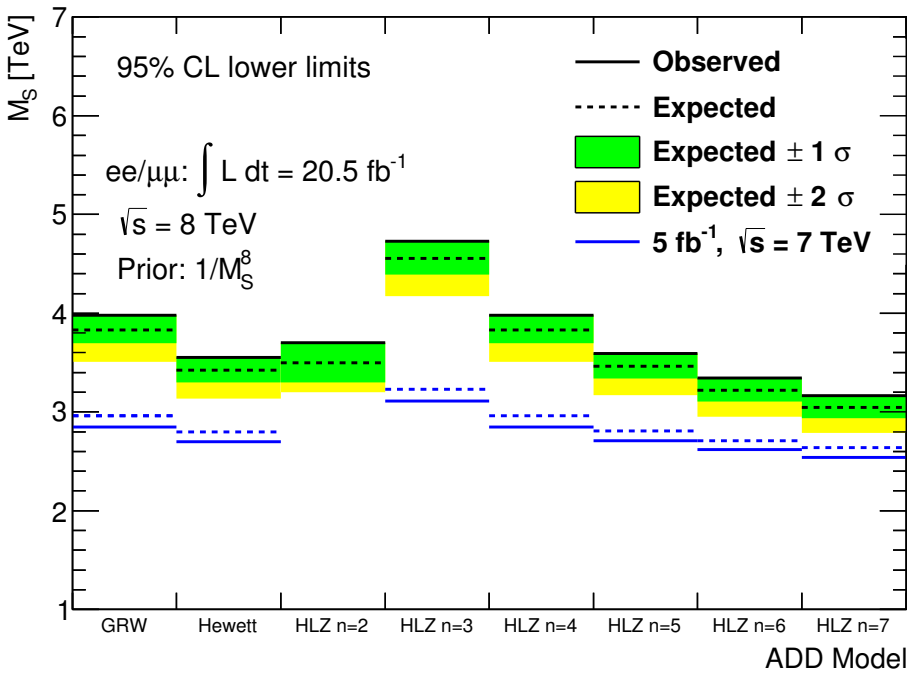


Figure 8.8: Observed and expected limit values on the ADD model searches in the dilepton channel using a uniform positive prior in $1/M_S^8$.

Table 8.3: Derived p-values for the CI LL, LR, RL and RR chiral structures and constructive and destructive interference models in dielectron and dilepton searches using uniform positive priors in $1/\Lambda^2$ and $1/\Lambda^4$.

p-value [%]	$1/\Lambda^2$		$1/\Lambda^4$	
	Constructive	Destructive	Constructive	Destructive
LL: ee	68	48	49	56
LR: ee	57	39	40	44
RL: ee	56	34	51	25
RR: ee	56	91	40	55
LL: $\ell\ell$	62	41	55	75
LR: $\ell\ell$	13	65	52	45
RL: $\ell\ell$	12	68	48	18
RR: $\ell\ell$	78	68	55	44

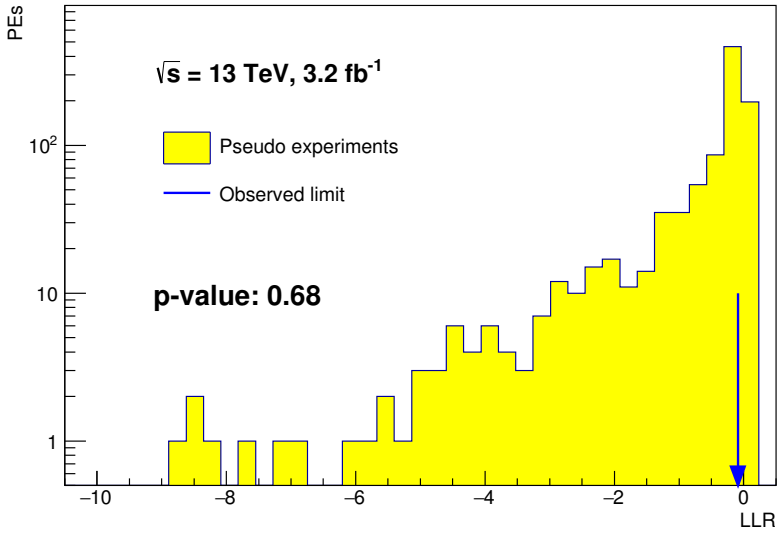


Figure 8.9: Negative log-likelihood distribution resulting from pseudo-experiments with fluctuations on the number of expected SM-only events in the dielectron channel for the CI LL constructive model using uniform positive priors in $1/\Lambda^2$.

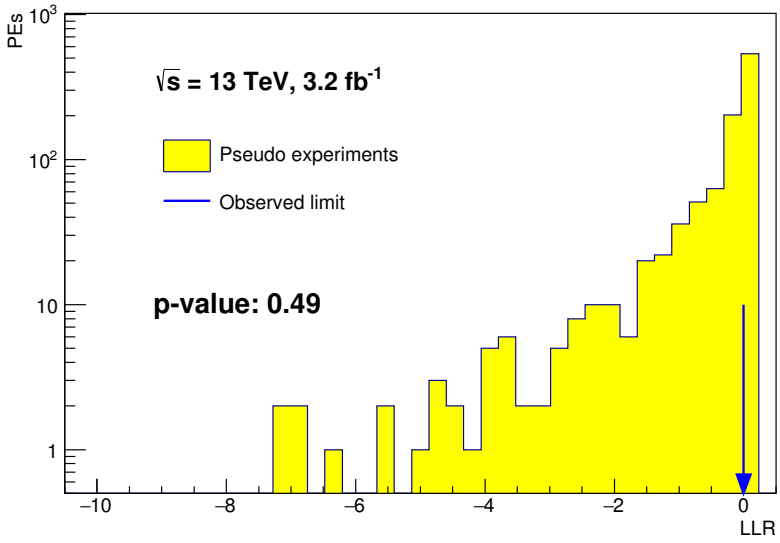


Figure 8.10: Negative log-likelihood distribution resulting from pseudo-experiments with fluctuations on the number of expected SM-only events in the dielectron channel for the CI LL constructive model using uniform positive priors in $1/\Lambda^4$.

8.2.2 Limit Setting

In an absence of signal, lower exclusion limits are set on the scale Λ. The posterior probability density distributions for data using the $1/\Lambda^2$ and $1/\Lambda^4$ priors in the electron

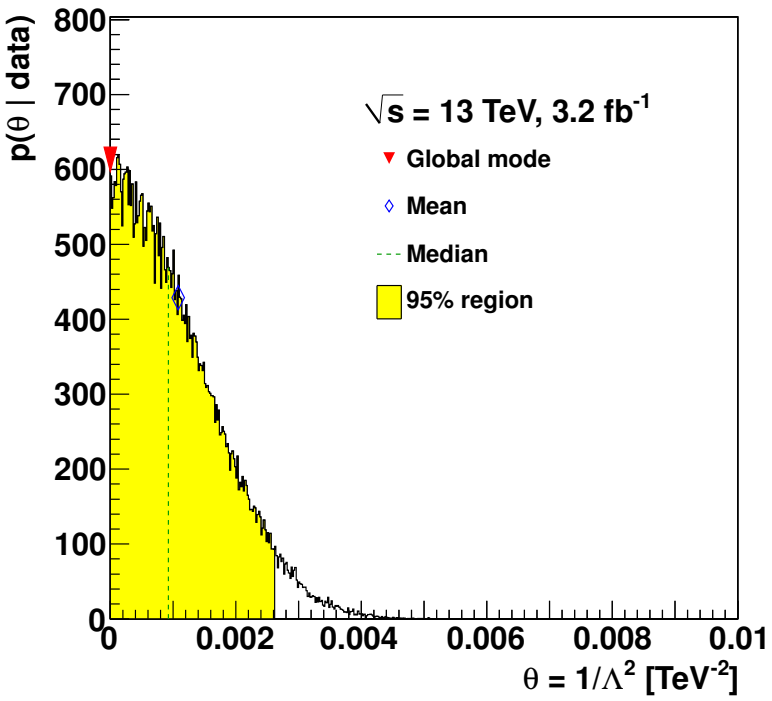


Figure 8.11: Posterior probability density distributions for the CI LL constructive model in the dielectron channel, with uniform positive priors in $1/\Lambda^2$.

channel are shown in Figures 8.11 and 8.12. The Θ_{Lim} value satisfying Equation 7.4 for these posterior distributions is the observed 95% C.L. on this model. The same procedure is conducted on 1000 PEs resulting in 1000 expected limits, the median of which is taken as the expected limit on the model. These limits are displayed in Figure 8.13 and Figure 8.14 for the electron channel, under both prior probabilities.

The results range from $\Lambda >$ 14.4–25.2 TeV, with the benchmark CI LL constructive model at $\Lambda >$ 19.5 TeV and $\Lambda >$ 25.2 TeV using flat prior probabilities in $1/\Lambda^2$ for the dielectron and dilepton final states, respectively. All results are presented in Table 8.4.

Figures 8.15, 8.16, 8.17 and 8.18 display the observed and expected limits for the dielectron and dilepton channels including the 1 and 2 sigma limit bands for both sets of prior probabilities. Included are the $\sqrt{s} = 8$ TeV results for comparison. The LR model from the $\sqrt{s} = 8$ TeV result is not included as this set LR = RL = 1, so is not a direct comparison to this result which separates the two models. The dataset used for the CI

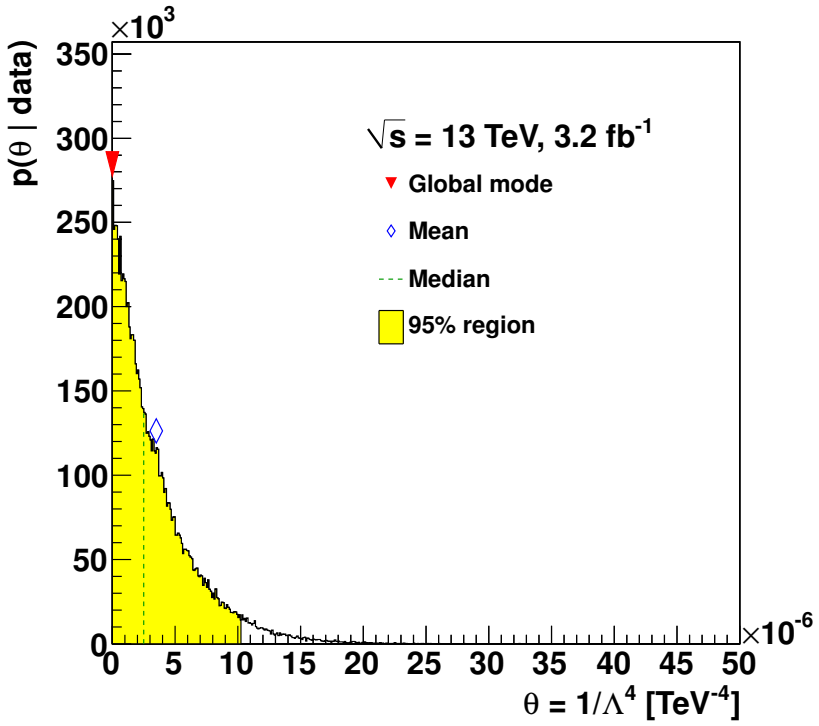


Figure 8.12: Posterior probability density distributions for the CI LL constructive model in the dielectron channel, with uniform positive priors in $1/\Lambda^4$.

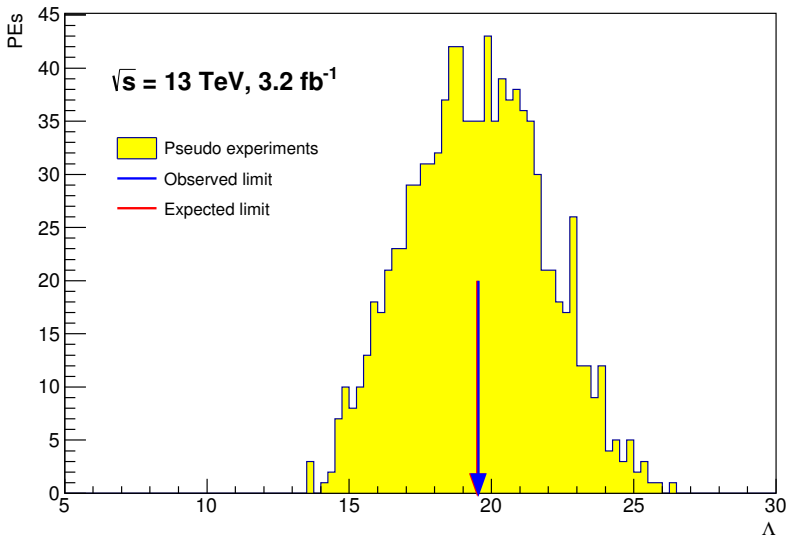


Figure 8.13: Distribution of observed and expected limit values in the dielectron channel for the CI LL constructive model with uniform positive priors flat in $1/\Lambda^2$.

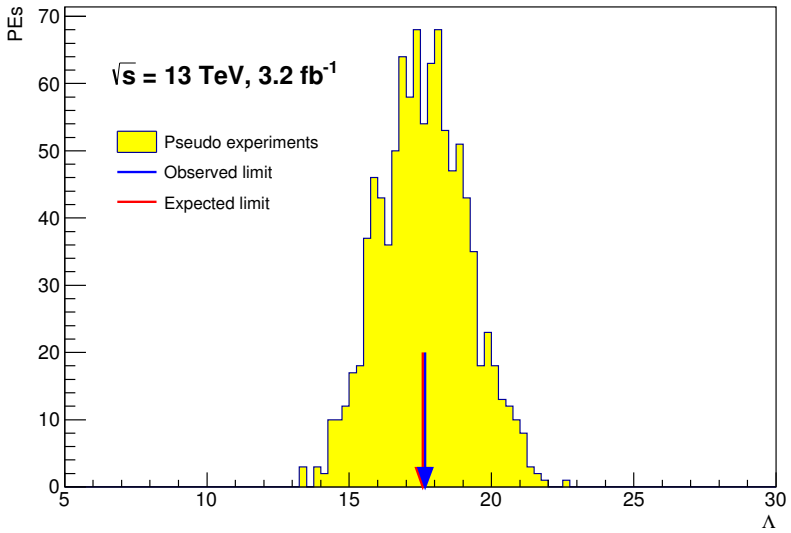


Figure 8.14: Distribution of observed and expected limit values in the dielectron channel for the CI LL constructive model with uniform positive priors flat in $1/\Lambda^4$.

Table 8.4: Observed and expected lower exclusion limits on Λ for CI LL, LR, RL and RR chiral structures and constructive and destructive interference models.

Expected and observed lower limits on Λ [TeV]									
Channel	Prior	Left-Left		Left-Right		Right-Left		Right-Right	
		Const.	Destr.	Const.	Destr.	Const.	Destr.	Const.	Destr.
Exp: ee	$1/\Lambda^2$	19.5	15.8	18.7	16.5	18.4	16.5	18.4	16.6
Obs: ee		19.5	15.5	18.7	16.2	18.5	16.4	18.5	16.4
Exp: ee	$1/\Lambda^4$	17.6	14.7	16.9	15.3	16.8	15.3	16.8	15.4
Obs: ee		17.7	14.4	17.0	15.0	16.8	15.1	16.8	15.1
Exp: $\ell\ell$	$1/\Lambda^2$	22.3	17.0	21.3	18.0	20.7	18.1	21.6	17.5
Obs: $\ell\ell$		25.2	17.8	24.1	19.2	23.5	19.6	24.6	18.2
Exp: $\ell\ell$	$1/\Lambda^4$	20.2	15.9	19.6	17.0	19.1	17.0	19.5	16.5
Obs: $\ell\ell$		22.2	16.7	21.3	17.8	21.0	18.1	21.7	17.0

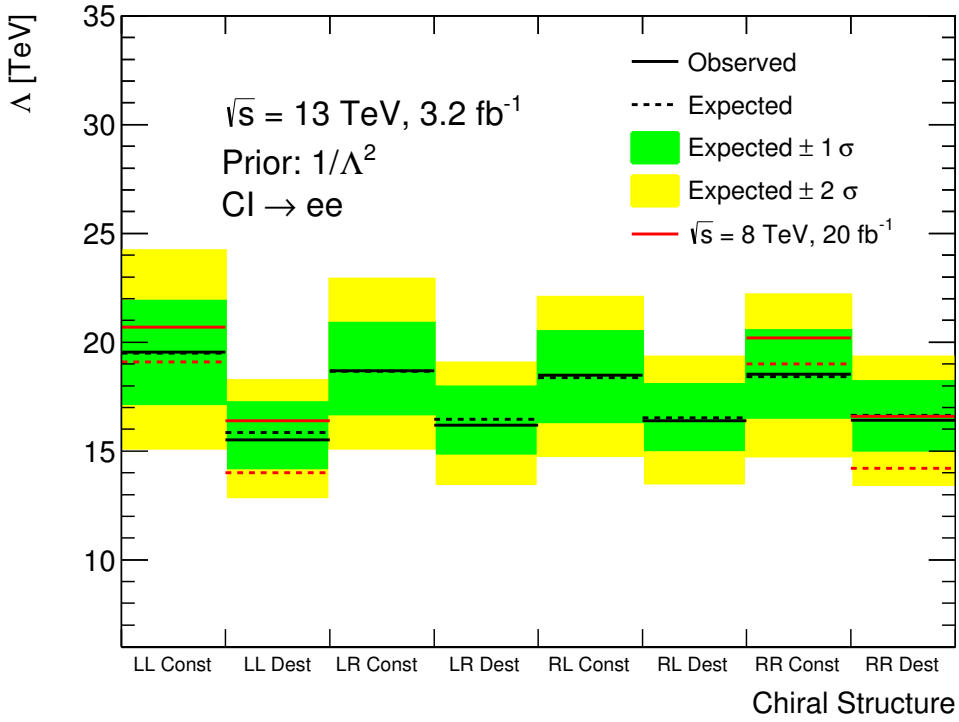


Figure 8.15: Observed and expected limit values on the CI model searches in the dielectron channel using a uniform positive prior in $1/\Lambda^2$.

search at $\sqrt{s} = 13$ TeV is much smaller than at $\sqrt{s} = 8$ TeV, but with the increase in cross-section the results are comparable. As more data is collected in 2016, these results will rapidly overtake the previous $\sqrt{s} = 8$ TeV limits.

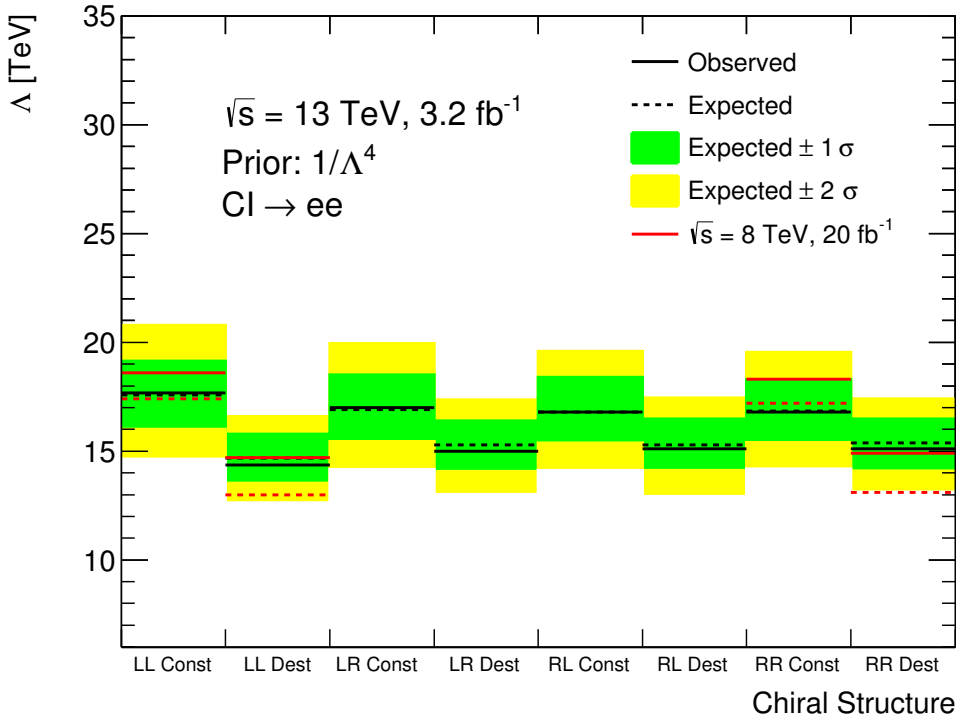


Figure 8.16: Observed and expected limit values on the CI model searches in the dielectron channel using a uniform positive prior in $1/\Lambda^4$.

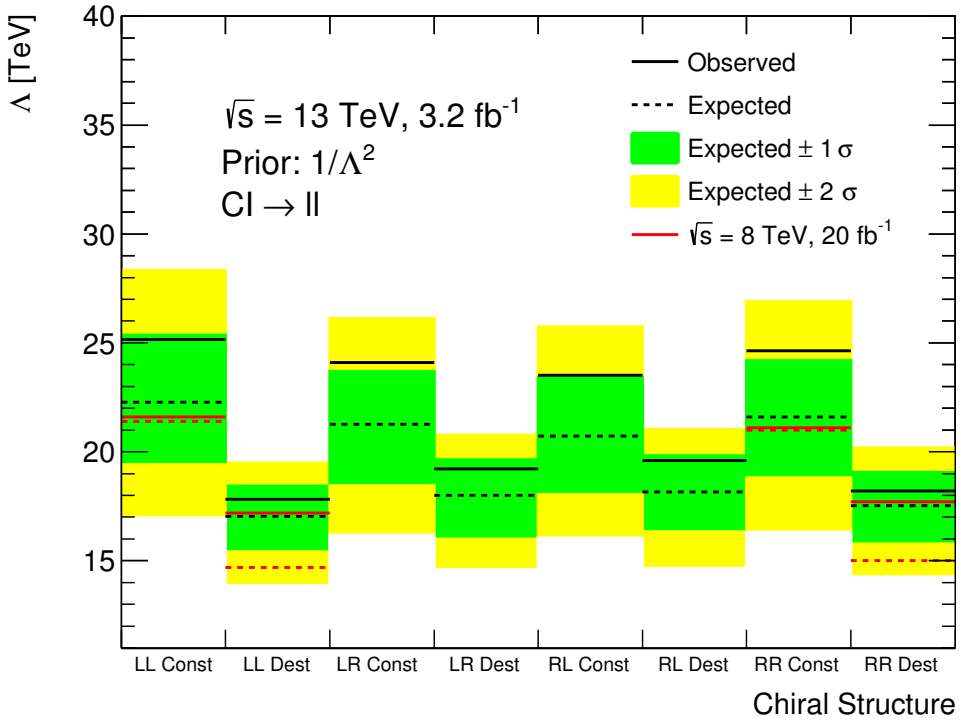


Figure 8.17: Observed and expected limit values on the CI model searches in the dilepton channel using a uniform positive prior in $1/\Lambda^2$.

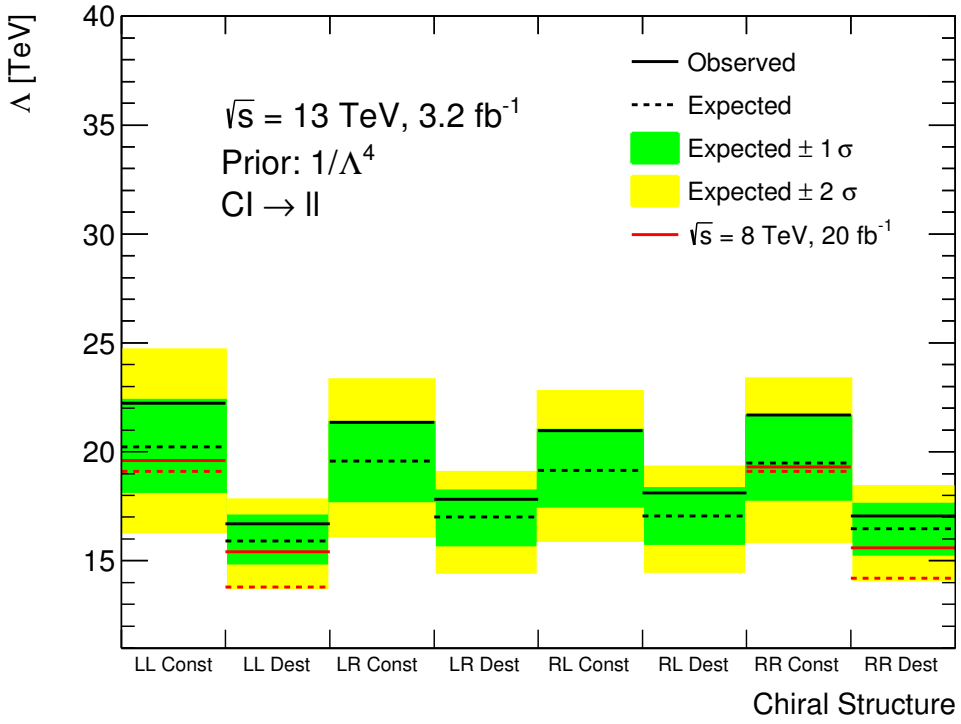


Figure 8.18: Observed and expected limit values on the CI model searches in the dilepton channel using a uniform positive prior in $1/\Lambda^4$.

"In the strict formulation of the law of causality - if we know the present, we can calculate the future - it is not the conclusion that is wrong but the premise."

– Werner Heisenberg

9 Conclusion

A search for non-resonant new phenomena in the dielectron channel has been carried out using both the 2012 and 2015 LHC proton-proton collision datasets. This corresponds to 20.3 fb^{-1} at $\sqrt{s} = 8$ TeV for the ADD large extra dimensions (LED) search and 3.2 fb^{-1} at $\sqrt{s} = 13$ TeV for the four fermion contact interaction (CI) search. This study builds upon previous ATLAS searches using the dielectron invariant mass as a search variable and combines with the dimuon channel for a dilepton result. No significant deviations from the Standard Model predictions are observed and lower limits are placed on the LED and CI scales.

In the $\sqrt{s} = 8$ TeV ADD search the most restrictive 95% C.L. lower limits are obtained in the combined dilepton channel, yielding $M_S > 5.0$ TeV for the HLZ $n = 3$ ADD model with a prior flat in $1/M_S^4$. This is an increase from $M_S > 3.88$ TeV from the ATLAS $\sqrt{s} = 7$ TeV dilepton analysis [30]. For the first time in an ATLAS search, limits on the HLZ $n = 2$ model are placed, the strongest of which are $M_S > 4.2$ TeV in the dilepton channel under a prior flat in $1/M_S^4$.

In the $\sqrt{s} = 13$ TeV CI search the most restrictive 95% C.L. lower limits are obtained in the combined dilepton channel, yielding $\Lambda > 25.2$ TeV for the CI LL constructive model with a prior flat in $1/\Lambda^2$. This is an increase from $\Lambda > 21.6$ TeV from the ATLAS $\sqrt{s} = 8$ TeV dilepton analysis [1]. Other recent analysis at $\sqrt{s} = 13$ TeV measured limits in the dijet final state search at $\Lambda > 17.5$ TeV in the same model [87].

The results presented in this thesis represent the most stringent measurements on LED and CI to date. With run 2 continuing in April 2016, more data will be collected and the dilepton invariant mass spectrum will remain one of the cleanest observables to search for new physics. Another method to increase sensitivity, employed successfully in the $\sqrt{s} = 8$ TeV CI analysis is the use of angular variables [1]. A similar investigation is planned for the full 2015 and 2016 combined datasets to increase the sensitivity to CI signals.

When the design centre-of-mass energy of $\sqrt{s} = 14$ TeV is reached, new physics models will gain a boost in cross-section higher than that of the SM background processes. The ability to probe these non-resonant models comes most readily from increases in the energy frontier rather than the gain in collected data from projects such as the high-luminosity LHC [88]. A study into the increase in limits for the ADD search at projected luminosities for $\sqrt{s} = 13$ (14) TeV is presented in Appendix D. With the increase in ADD limits the distance between the electroweak and the fundamental Planck scales grows — the problem the model aimed to address originally. As such the model is becoming out of favour.

While the proposed models investigated in this thesis may not currently be evident in nature, the SM is known not to be an ultimate theory, and an extension must be included to explain the measured experimental results. The four-fermion contact interaction model allows searches for new physics at energies much higher than achievable for direct detection and will remain of high priority in the ATLAS exotics group.

Appendices

A

Kinematic Plots

The dielectron invariant mass spectrum for the HLZ $n = 2$ is displayed in Figure A.1. This shows the full data selection from 2012, with expected SM process and new physics compared. The ratio plot in the inset shows excellent agreement between data and background and no deviation from the SM.

Figures A.2, A.3 and A.4 show the $\sqrt{s} = 8$ TeV analysis E_T, η and ϕ distributions, respectively. Similarly the results from the $\sqrt{s} = 13$ TeV analysis are shown in A.5, A.6 and A.7 for the E_T, η and ϕ distributions, respectively.

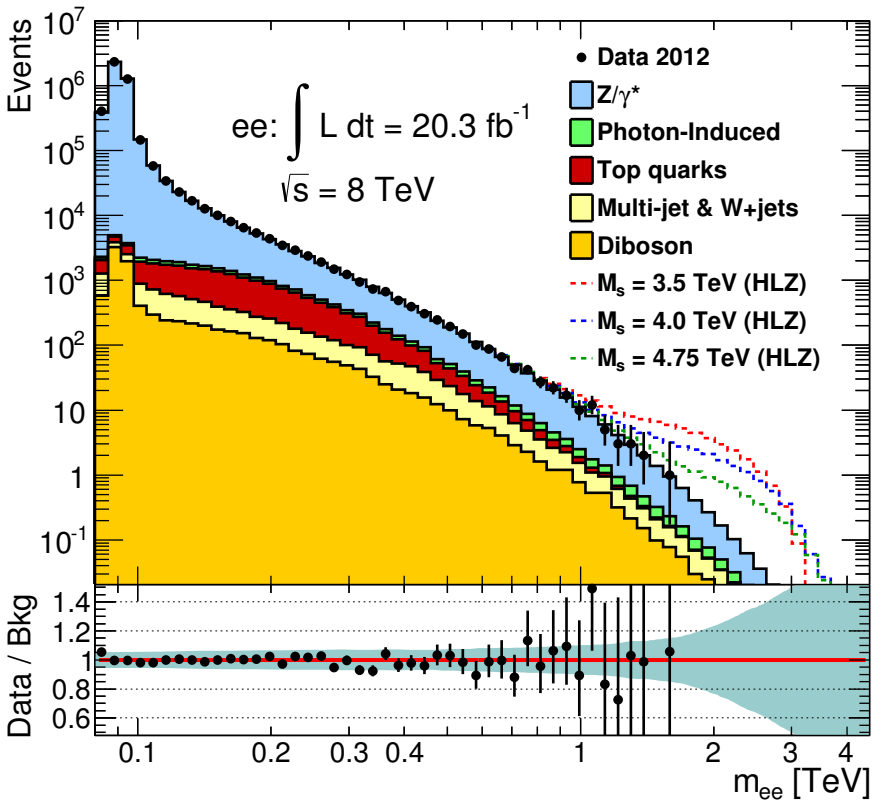


Figure A.1: Invariant mass spectrum of data and background processes with ADD HLZ ($n = 2$) signal overlaid, the bottom inset shows the ratio between data and background.

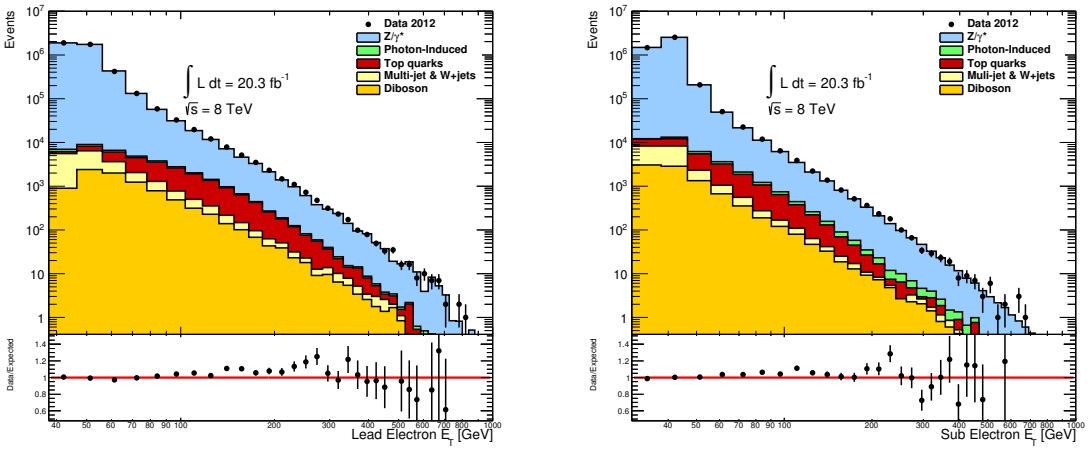


Figure A.2: $\sqrt{s} = 8$ TeV E_T distribution for separated leading and subleading electrons.

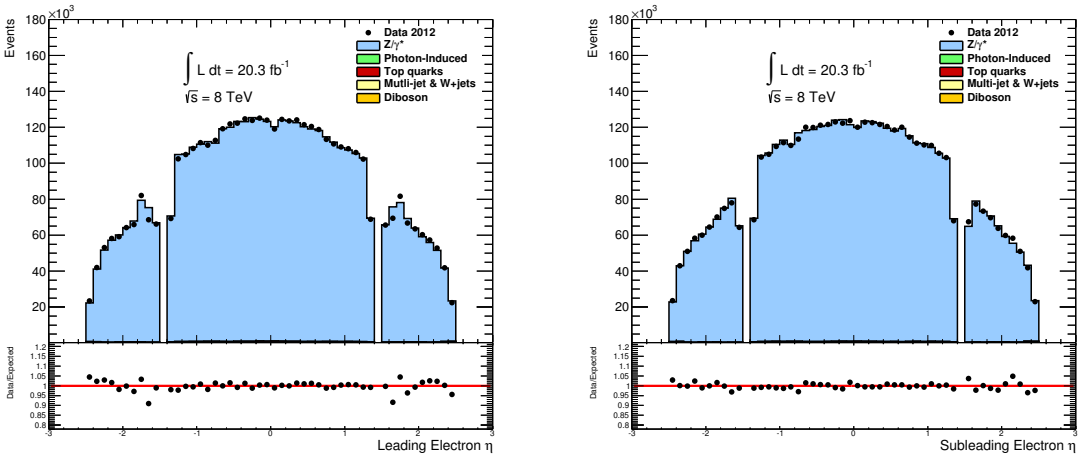


Figure A.3: $\sqrt{s} = 8$ TeV η distribution for separated leading and subleading electrons.

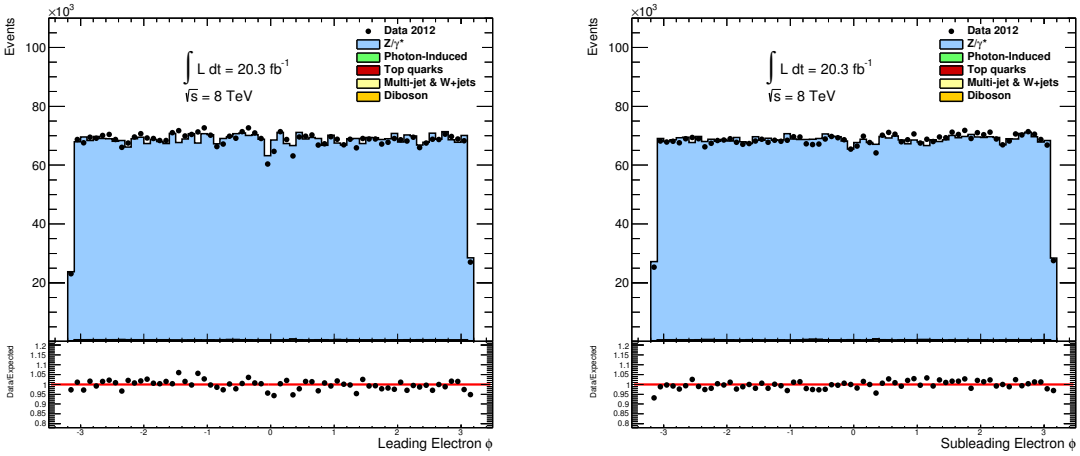


Figure A.4: $\sqrt{s} = 8$ TeV ϕ distribution for separated leading and subleading electrons.

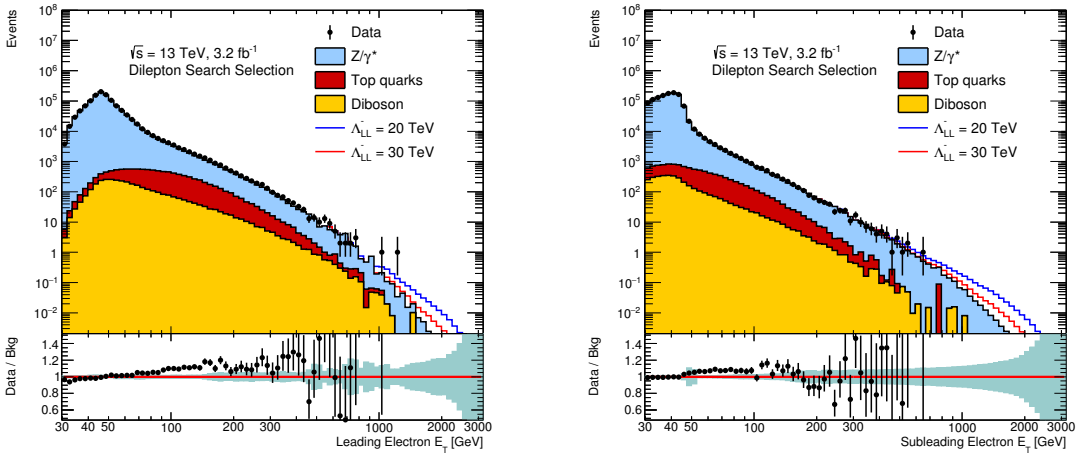


Figure A.5: $\sqrt{s} = 13$ TeV E_T distribution for separated leading and subleading electrons.

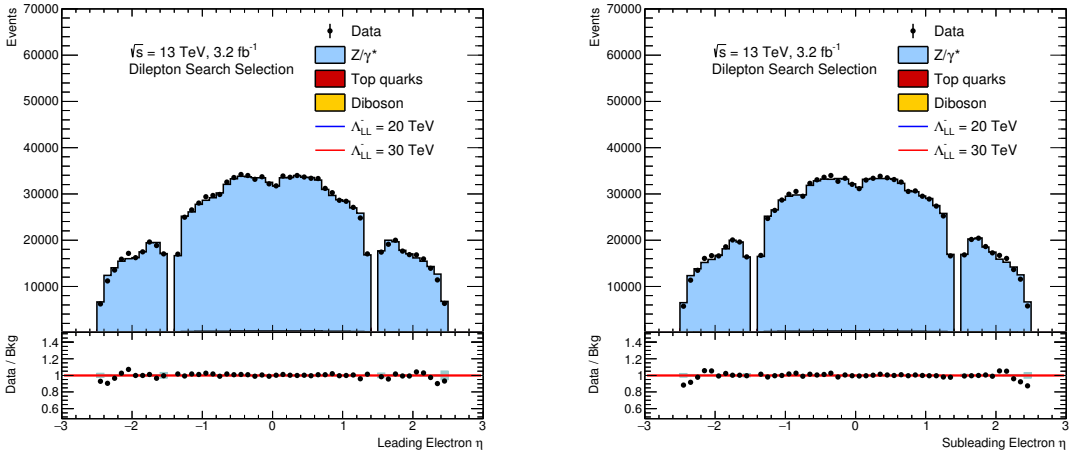


Figure A.6: $\sqrt{s} = 13$ TeV η distribution for separated leading and subleading electrons.

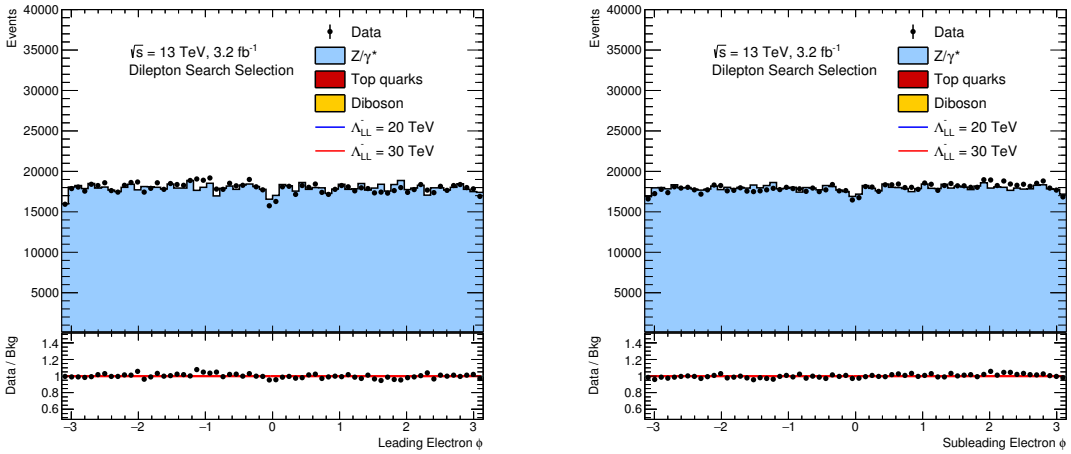


Figure A.7: $\sqrt{s} = 13$ TeV ϕ distribution for separated leading and subleading electrons.

B

Dielectron p_T Re-weighting

A cross-check study was performed to ascertain the affect of the known dielectron p_T disagreement seen between MC and data on the invariant mass spectrum. The dielectron p_T ratio between data and MC was extracted from Figure B.1, and was used to re-weight MC on an event by event bases according to its dielectron p_T. The corresponding DY nominal/p_T re-weighted invariant mass distributions were compared, as shown in Figure B.2. In the region where data is available for this analysis, the effect is negligible. At high-mass the affect grows to 2.5% which is still considered small and is well within the systematic uncertainties of this search.

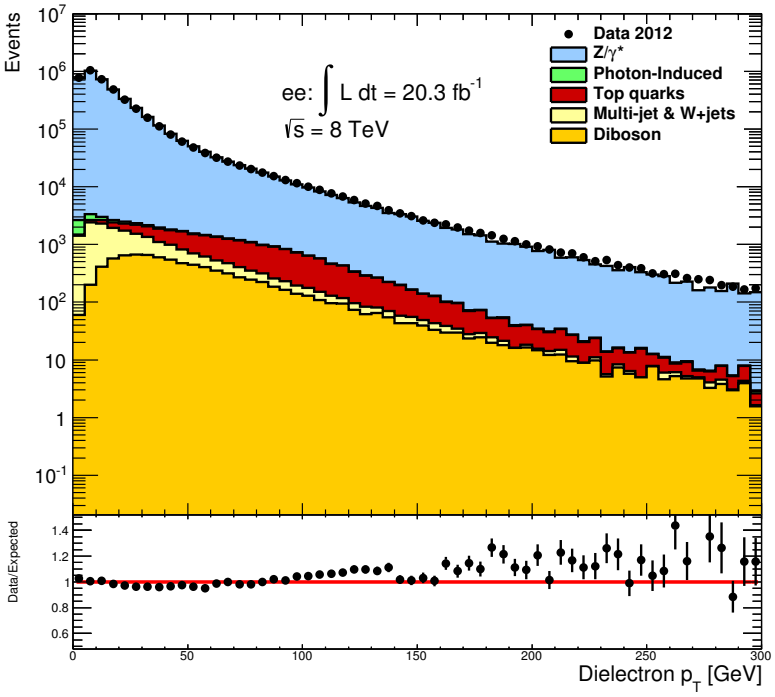


Figure B.1: Dielectron p_T distribution.

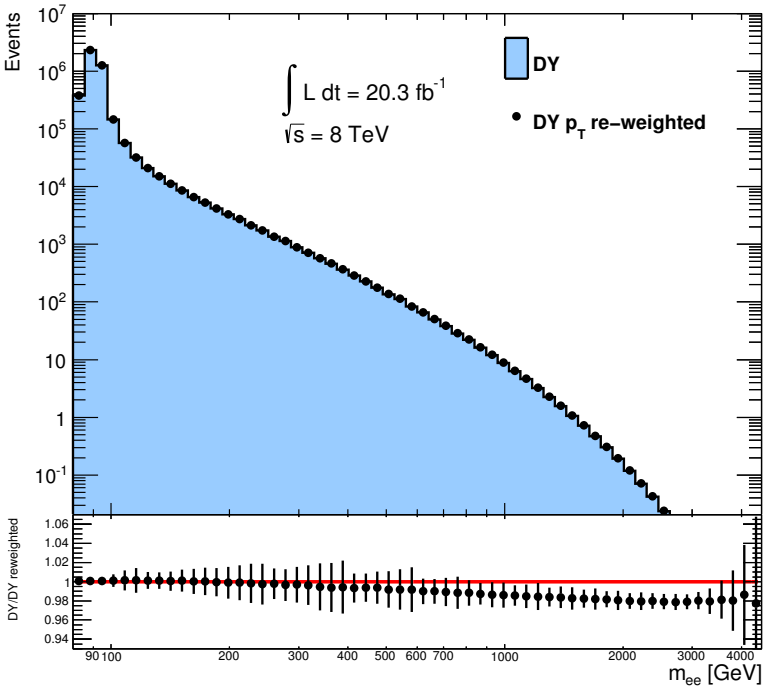


Figure B.2: p_T re-weighted DY to data invariant mass comparison.

C

Signal Parameterisations

The additional CI LL destructive signal parameterisations are displayed in Figure C.1. The signal systematic parameterisations are shown in Figures C.2, C.3, C.4 and C.5 for the constructive interference model and Figures C.6, C.7, C.8 and C.9 for the destructive interference model.

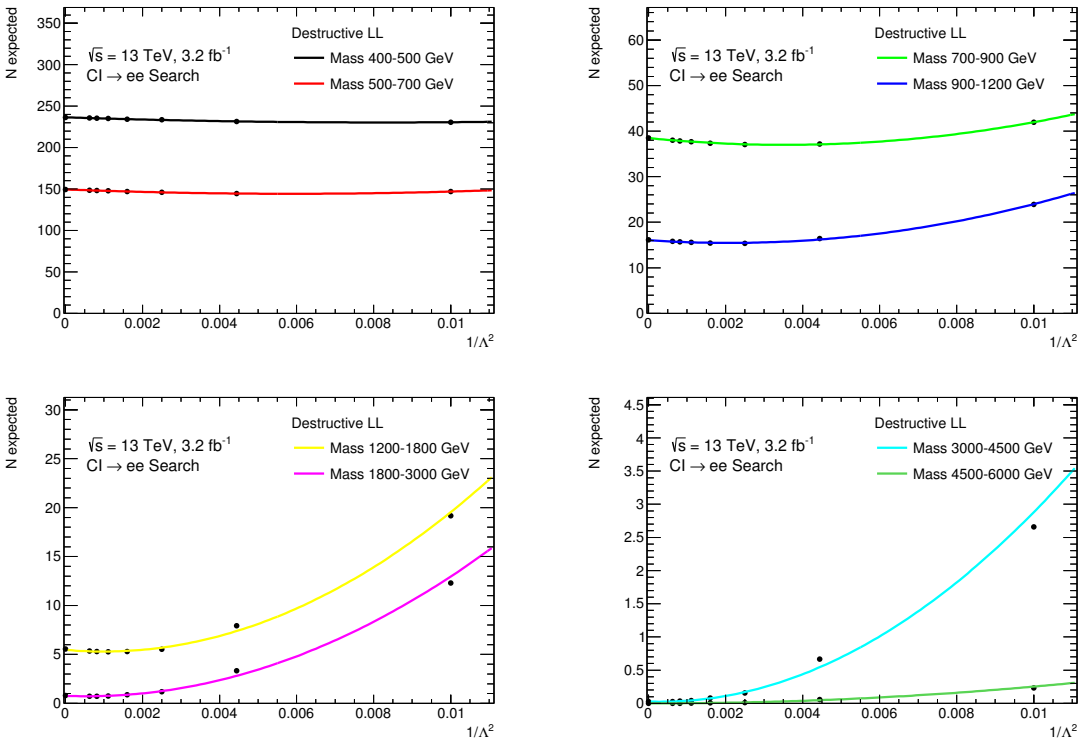


Figure C.1: Number of expected events in the electron channel for the CI LL destructive interference search regions as a function of $1/\Lambda^2$.

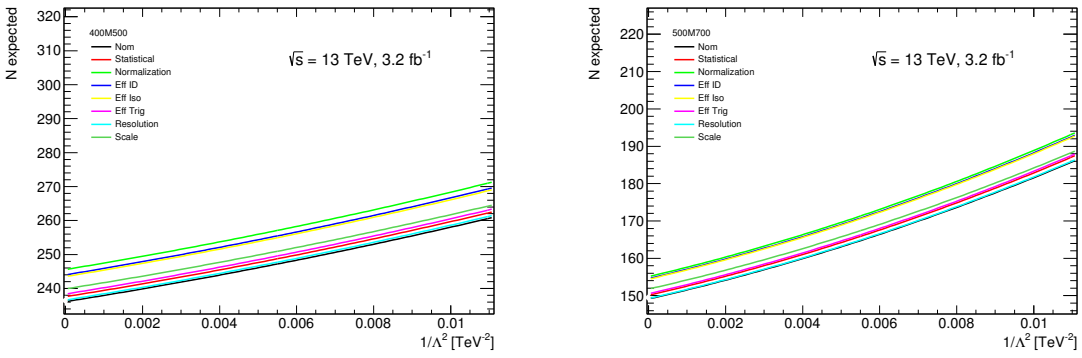


Figure C.2: Constructive LL signal systematic parameterisations of number of expected events in mass region 400–500 GeV (left) and 500–700 GeV (right).

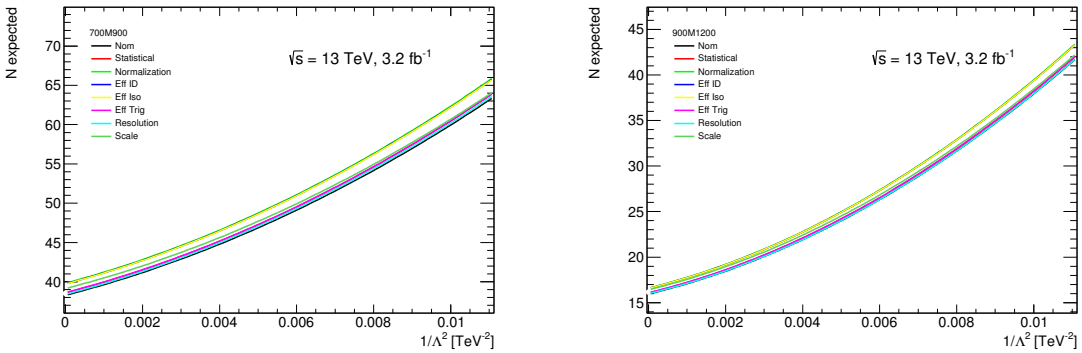


Figure C.3: Constructive LL signal systematic parameterisations of number of expected events in mass region 700–900 GeV (left) and 900–1200 GeV (right).

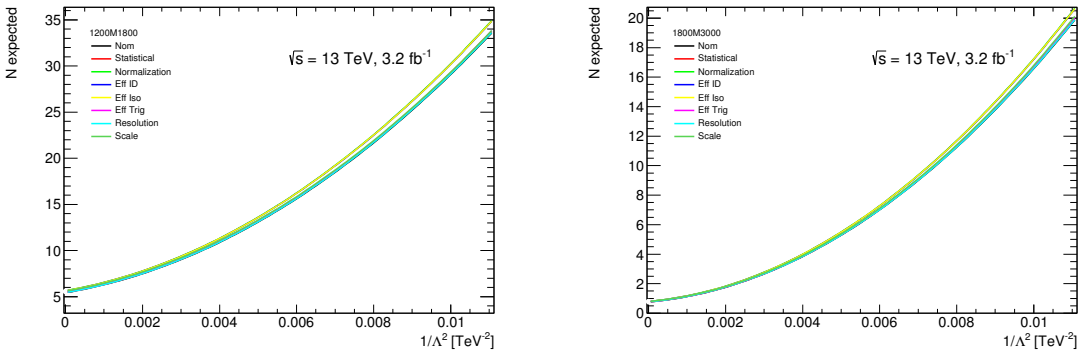


Figure C.4: Constructive LL signal systematic parameterisations of number of expected events in mass region 1200–1800 GeV (left) and 1800–3000 GeV (right).

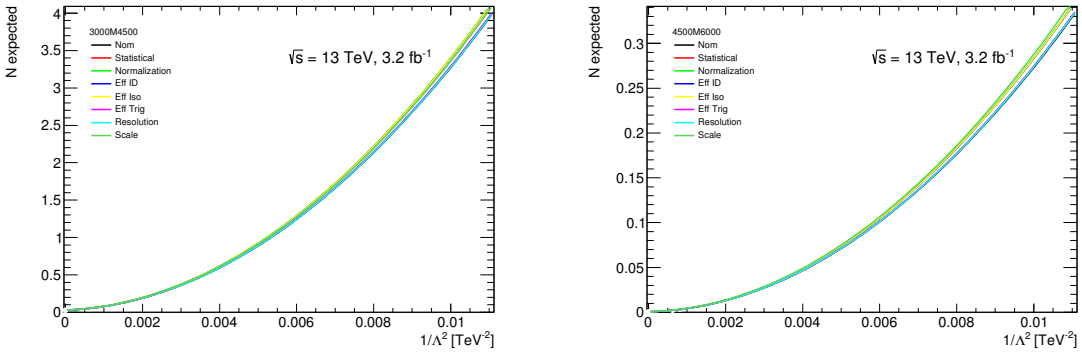


Figure C.5: Constructive LL signal systematic parameterisations of number of expected events in mass region 3000–4500 GeV (left) and 4500–6000 GeV (right).

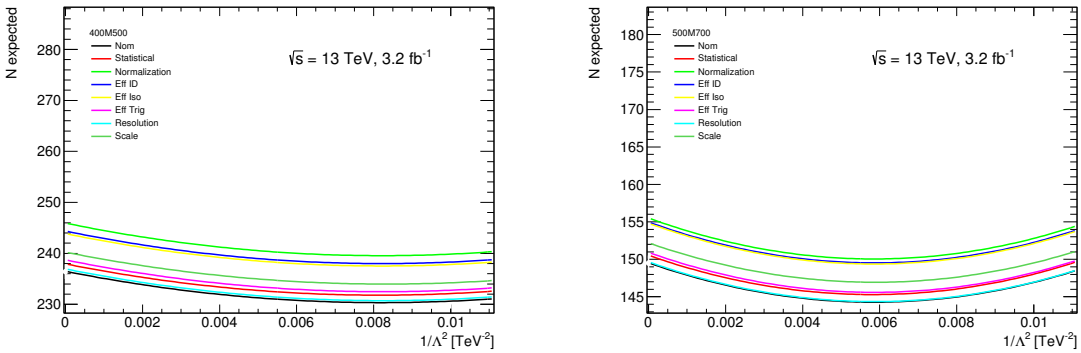


Figure C.6: Destructive LL signal systematic parameterisations of number of expected events in mass region 400–500 GeV (left) and 500–700 GeV (right).

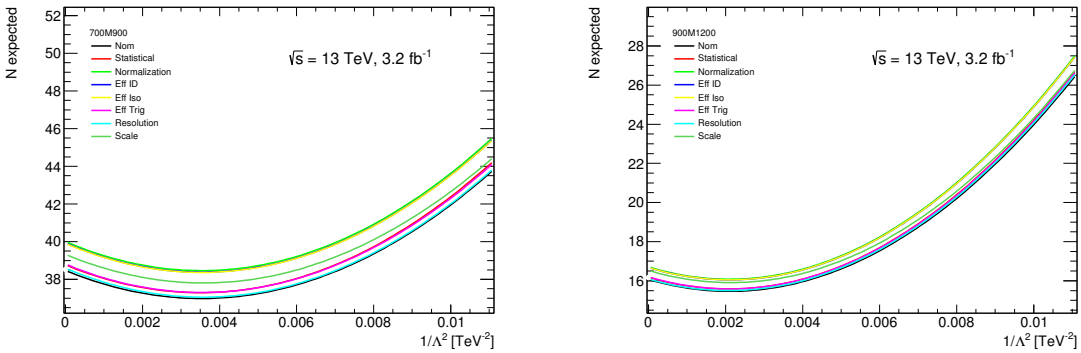


Figure C.7: Destructive LL signal systematic parameterisations of number of expected events in mass region 700–900 GeV (left) and 900–1200 GeV (right).

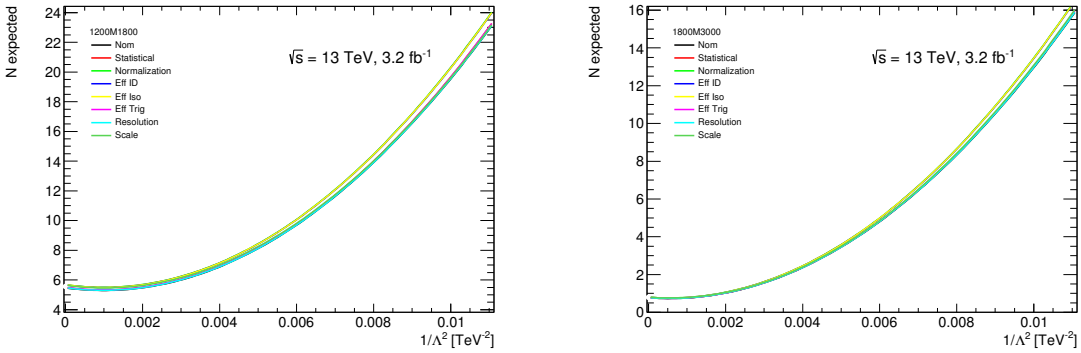


Figure C.8: Destructive LL signal systematic parameterisations of number of expected events in mass region 1200–1800 GeV (left) and 1800–3000 GeV (right).

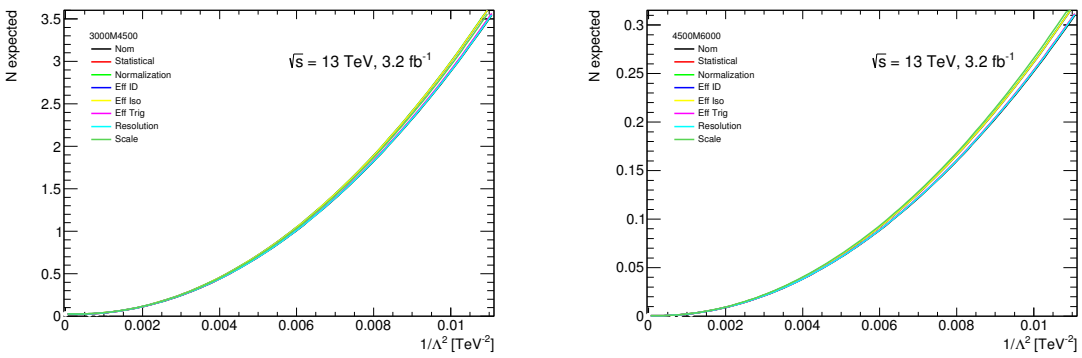


Figure C.9: Destructive LL signal systematic parameterisations of number of expected events in mass region 3000–4500 GeV (left) and 4500–6000 GeV (right).

D

ADD Projections

This appendix describes the projected lower limits on the ADD model at higher centre-of-mass energies and integrated luminosities. This is performed at $\sqrt{s} = 13$ and 14 TeV with a range of integrated luminosities using the available $\sqrt{s} = 8$ TeV ADD signal and NNLO DY background samples. Whilst the ADD search bin was optimised (explained in Section 7.4.1) for the $\sqrt{s} = 8$ TeV analysis, the same search bin is used in these projection studies. The MC event generator programme SHERPA is used to calculate the cross-sections in the truth invariant mass range of 1.9–4.5 TeV at $\sqrt{s} = 8$, 13 and 14 TeV for each of the M_S values generated for the $\sqrt{s} = 8$ TeV search. This includes the essentially DY only $M_S = 50$ TeV sample. The cross-section ratio between the projected centre-of-mass energy and the $\sqrt{s} = 8$ TeV is used as a scale factor to estimate the number of reconstructed events at a given collision energy and integrated luminosity. The cross-sections and scale factors are shown in Table D.1.

Table D.1: Table presenting the cross-section scale factors used in ADD $\sqrt{s} = 13$ and 14 projections.

M_S [TeV]	$\sigma_{8\text{ TeV}}$ [fb]	$\sigma_{13\text{ TeV}}$ [fb]	$\sigma_{14\text{ TeV}}$ [fb]	$\sigma_{13\text{ TeV}}/\sigma_{8\text{ TeV}}$	$\sigma_{14\text{ TeV}}/\sigma_{8\text{ TeV}}$
3.0	6.8	54.7	61.6	8.04	9.06
3.25	4.2	35.7	47.6	8.53	11.37
3.5	2.7	24.5	33.0	8.97	12.10
3.75	1.8	17.0	22.9	9.22	12.44
4.0	1.4	12.4	16.6	9.12	12.14
4.75	0.7	5.6	7.4	7.63	10.12
50.0	0.5	2.3	2.8	4.48	5.34

Each $\sqrt{s} = 8$ TeV expected M_S signal yield in the search bin is scaled by the appropriate cross-section scale factor to estimate the number of expected events. A parabola is fit to these estimates as with the $\sqrt{s} = 8$ TeV analysis to create a parameterisation of the number of expected events as a function of $1/M_S^4$. The SM estimate used in these studies is the nominal NNLO DY estimate only, scaled using the $M_S = 50$ TeV scale factor. The dominant systematic uncertainty on PDF variation is used in these studies, firstly as the nominal $\sqrt{s} = 8$ TeV uncertainty, then with the uncertainty halved. As at higher centre-of-mass energies the PDFs will be better constrained for a given invariant mass, the full and halved uncertainties are to give a sense of where between the limits will lie. These inputs are propagated into the limit setting framework, and using 1000 SM only pseudo experiments an expected limit can be obtained for a given integrated luminosity. The results of which are displayed in Figure D.1 under a prior flat in $1/M_S^8$.

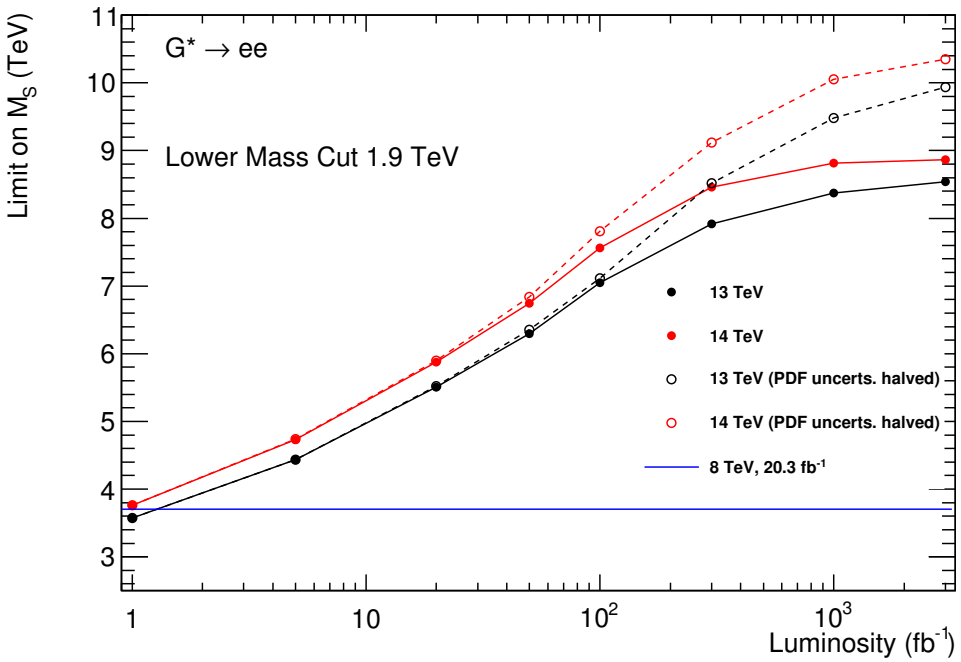


Figure D.1: ADD $\sqrt{s} = 13$ and 14 lower limit projections. Full and half $\sqrt{s} = 8$ TeV PDF variation systematic uncertainties are used for comparison.

With a jump in centre-of-mass energy from $\sqrt{s} = 8$ TeV to $\sqrt{s} = 13$ TeV the limits increase drastically, requiring only 1-2 fb^{-1} to surpass the $\sqrt{s} = 8$ TeV result. A modest

boost in sensitivity is also seen with an increase to $\sqrt{s} = 14$ TeV. As the collected data increases, the limits increase relatively slowly, with a factor of 10 increase in dataset size the limits increase by only $\sim$ 1-2 TeV. After $\sim$ 100 fb^{-1} (depending on the value of the systematic uncertainty) the limits begin to plateau, with a gain in data not increasing sensitivity to the ADD model.

Bibliography

[1] ATLAS Collaboration, *Search for contact interactions and large extra dimensions in the dilepton channel using proton-proton collisions at* $\sqrt{s} = 8$ TeV *with the ATLAS detector*, Eur. Phys. J. C74.12 3134 (2014).

[2] T. Varol, *Search for contact interactions and large extra dimensions in dimuon events from pp collisions at* $\sqrt{s} = 8$ TeV *with the ATLAS detector*, Ph.D. thesis, University of Massachusetts Amherst (2015).

[3] M. C. Bret, *Search for Quantum Black Holes and ADD Extra Dimensions in the opposite sign dimuon channel in proton–proton collisions with the ATLAS detector at* $\sqrt{s} = 8$ TeV, Ph.D. thesis, Queen Mary University of London (2015).

[4] ATLAS Collaboration, *Search for new phenomena in the dilepton final state using proton-proton collisions at* $\sqrt{s} = 13$ TeV *with the ATLAS detector*, ATLAS-CONF-2015-070.

[5] J. Beringer et. al. (Particle Data Group), *Review of Particle Physics*, Chin. Phys. C 38 (2014) 090001.

[6] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys.22 (1961) 579-588.

[7] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc.C680519 (1968) 367-377.

[8] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett.19 (1967) 1264-1266.

[9] F. Englert, R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. 13 (1964) 321-323.

[10] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett.13 (1964) 508-509.

[11] G. S. Guralnik, C. R. Hagen, T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. 13 (1964) 585-587.

[12] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B716 (2012) 1-29.

[13] CMS Collaboration, *Observation of a new boson at a mass of 125* GeV*with the CMS experiment at the LHC*, Phys. Lett. B716 (2012) 30-61.

[14] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys. Lett. B429 (1998) 263-272.

[15] V. A. Rubakov, *Large and infinite extra dimensions: An Introduction*, Phys. Usp. 44 (2001) 871–893.

[16] T.Gleisberg et. al., *Helicity formalism for spin-2 particles*, JHEP 09 (2003) 001.

[17] G. F. Giudice, R. Rattazzi, J. D. Wells, *Quantum gravity and extra dimensions at high-energy colliders*, Nucl. Phys. B544 (1999) 3-38.

[18] J. L. Hewett, *Indirect collider signals for extra dimensions*, Phys. Rev. Lett. 82 (1999) 4765–4768.

[19] T. Han, J. D. Lykken, R.-J. Zhang, *On Kaluza-Klein states from large extra dimensions*, Phys. Rev. D59 (1999) 105006.

[20] I. DSouza, C. S.Kalman, *Preons; Models of leptons, quarks and gauge bosons as composite objects*, (World Scientific, Singapore) 1992.

[21] E. Eichten, K. D. Lane, M. E. Pes, *New tests for quark and lepton substructure*, Phys. Rev. Lett. 50, (1983) 811 .

[22] L3 Collaboration, *Search for extra dimensions in boson and fermion pair production in e^+e^- interactions at LEP*, Phys. Lett. B 470, (1999) 281.

[23] L3 Collaboration, *Search for low scale gravity effects in e^+e^- collisions at LEP*, Phys. Lett. B 464, (1999) 135.

[24] DELPHI Collaboration, *Measurement and interpretation of fermion-pair production at LEP energies of 183 and 189* GeV, Phys. Lett. B 485, (2000) 45.

[25] DELPHI Collaboration, *Determination of the* $e^+e^- \to \gamma\gamma(\gamma)$ *cross-section at centre-of-mass energies ranging from 189* GeV *to 202* GeV, Phys. Lett. B 491, (2000) 67.

[26] OPAL Collaboration, *Multi-photon production in* e^+e^- *collisions at* $\sqrt{s} = 181 - 209$ GeV, Eur. Phys. J. C 26, (2003) 331.

[27] H1 Collaboration, *Search for compositeness, leptoquarks and large extra dimensions in eq contact interactions at HERA*, Phys. Lett. B 479, (2000) 358.

[28] D0 Collaboration, *Search for large extra spatial dimensions in the dielectron and diphoton channels in pp collisions at* $\sqrt{s} = 1.96$ TeV, Rev. Lett. 102, (2009) 051601.

[29] D0 Collaboration, *Search for large extra spatial dimensions in dimuon production with the D0 detector*, Phys. Rev. Lett. 95 (2005) 161602.

[30] ATLAS Collaboration, *Search for contact interactions and large extra dimensions in dilepton events from pp collisions at* $\sqrt{s} = 7$ TeV *with the ATLAS detector*, Phys. Rev. D 87 (2013) 015010.

[31] ATLAS Collaboration, *Search for Extra Dimensions in the Diphoton Final State in* $\sqrt{s} = 7$ TeV *pp collisions with 4.9* fb^{-1}, Phys. Lett. B710 (2012) 538-556.

[32] CMS Collaboration, *Search for large extra dimensions in dimuon and dielectron events in pp collisions at* $\sqrt{s} = 7$ TeV, Phys. Lett. B711 (2012). 15–34.

[33] CMS Collaboration, *Search for Signatures of Extra Dimensions in the Diphoton Mass Spectrum at the Large Hadron Collider*, Phys. Rev. Lett. (2012) 108.

[34] CMS Collaboration, *Search for physics beyond the standard model in dilepton mass spectra in proton-proton collisions at* $\sqrt{s} = 8$ TeV, JHEP 04 (2015) 025.

[35] ALEPH Collabaration, *Fermion pair production in* e^+e^- *collisions at 189–209–* GeV *and constraints on physics beyond the standard model*, Eur. Phys. J. C 49 (2007) 411-437.

[36] DELPHI Collabaration, *A Study of* $\text{b}\bar{\text{b}}$ *Production in* e^+e^- *Collisions at* $\sqrt{s} = 130 - \text{GeV} - 207 - \text{GeV}$, Eur. Phys. J. C 60 (2009) 1-15 .

[37] OPAL Collaboration, *Tests of the standard model and con- straints on new physics from measurements of fermion-pair production at 189 GeV at LEP*, Eur. Phys. J. C 13, (2000) 553.

[38] H1 Collabaration, *Search for Contact Interactions in $e^{\pm}p$ Collisions at HERA*, Phys. Lett. B 705 (2011) 52-58.

[39] D0 Collabaration, *Measurement of dijet angular distributions at* $\sqrt{s} = 1.96 - \text{TeV}$ *and searches for quark compositeness and extra spatial dimensions*, , Phys. Rev. Lett. 103 (2009) 191803.

[40] O. S. Brning et al., *LHC Design Report*, CERN.

[41] C. De Melis, *The CERN accelerator complex. Complexe des acclrateurs du CERN*, OPEN-PHO-ACCEL-2016-001.

[42] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST3 S08003 (2008).

[43] M.Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013.

[44] ATLAS Collaboration, twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[45] ATLAS Collaboration, twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[46] W.Lampl *et al.*, *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002.

[47] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, ATLAS-CONF-2014-032.

[48] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047.

[49] ATLAS Collaboration, *Electron identification measurements in ATLAS using* $\sqrt{s} =$ *13* TeV *data with 50 ns bunch spacing*, ATL-PHYS-PUB-2015-041.

[50] ATLAS Collaboration, *Performance of Particle Identification with the ATLAS Transition Radiation Tracker*, Particles and fields. Proceedings (2011).

[51] M. Aleksa et al., *ATLAS Liquid Argon Calorimeter Phase-I Upgrade Technical Design Report*, CERN-LHCC-2013-017.

[52] K. Rehermann, B. Tweedie, *Efficient Identification of Boosted Semileptonic Top Quarks at the LHC*, JHEP 03 (2011) 059.

[53] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A 506 (2003) 250–303.

[54] ATLAS Collaboration, *Measurement of the high-mass Drell-Yan differential cross-section in pp collisions at* $\sqrt{s} = 7$ TeV *with the ATLAS detector*, Phys. Lett. B725 (2013) 223–242.

[55] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C74 (10) (2014) 3071.

[56] ATLAS Collaboration, *Electron efficiency measurements using the 2015 LHC proton-proton collision data*, ATLAS-COM-CONF-2016-028.

[57] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, J. High Energy Phys. 1006 (2010) 043.

[58] H-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D 82 (2010) 074024.

[59] T. Sjostrand, S. Mrenna, P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. 178 (2008) 852–867.

[60] Y. Li, F. Petriello, *Combining QCD and electroweak corrections to dilepton production in FEWZ*, Phys. Rev. D 86 (2012) 094034.

[61] A.D Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. C 63 (2009) 189–285.

[62] S. Dulat et al., *The CT14 Global Analysis of Quantum Chromodynamics*, Phys. Rev. D93 (3) (2016) 033006.

[63] A.D Martin et al., *Parton distributions incorporating QED contributions*, Eur. Phys. J. C 39 (2005) 155–161.

[64] S. Frixione, B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, J. High Energy Phys. 0206 (2002) 029.

[65] J. Butterworth, J. R. Forshaw, M. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. C 72 (1996) 637–646.

[66] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, J. High Energy Phys. 0101 (2001) 010.

[67] T. Sjostrand, S. Mrenna, P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, J. High Energy Phys. 0605 (2006) 026.

[68] M. Czakon, A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. 185 (2014) 2930.

[69] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, J. High Energy Phys. 0207 (2002) 012.

[70] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP 0902 (2009) 007.

[71] ATLAS Collaboration, *Background studies for top-pair production in lepton plus jets final states in* $\sqrt{s} = 7TeV$ *ATLAS data*, ATLAS-CONF-2010-087.

[72] R. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B867 (2013) 244–289.

[73] J. Leveque (on behalf of the ATLAS Collaboration), *ATLAS Run 2 - Higgs Prospects*, ATL-PHYS-SLIDE-2015-450.

[74] M. L. Mangano, J. Rojo, *Cross Section Ratios between different CM energies at the LHC: opportunities for precision measurements and BSM sensitivity*, JHEP 08 (2012) 010.

[75] ATLAS Collaboration, *b-tagging in dense environments*, ATL-PHYS-PUB-2014-014.

[76] ATLAS Collaboration, *Measurement of the transverse momentum distribution of Z/γ^* bosons in protonproton collisions at $\sqrt{s}$=7 TeV with the ATLAS detector*, Phys. Lett. B705 (2011) 415–434.

[77] C. Anastasiou, L. J. Dixon, K. Melnikov, F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. D 69 (2004) 094008.

[78] ATLAS Collaboration, *Search for High-Mass Dilepton Resonances in pp Collisions at $\sqrt{s}$ = 8 TeV with the ATLAS Detector*, Phys. Rev. D90 (5) (2014) 052005.

[79] S. Alekhin, J. Blumlein, S. Moch, *The ABM parton distributions tuned to LHC data*, Phys. Rev. D89 (5) (2014) 054028.

[80] A. Cooper-Sarkar, *PDF Fits at HERA*, PoS EPS-HEP2011 (2011) 320.

[81] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G43 (2016) 023001.

[82] P. Motylinski et al., *Updates of PDFs for the 2nd LHC run*, in: International Conference on High Energy Physics 2014 (ICHEP 2014) Valencia, Spain, July 2-9.

[83] P. Jimenez-Delgado, E. Reya, *Delineating parton distributions and the strong coupling*, Phys. Rev. D89 (7) (2014) 074049.

[84] D. Bardin et al., *SANC integrator in the progress: QCD and EW contributions*, JETP Lett. 96 (2012) 285–289.

[85] J. Wenninger, *Energy Calibration of the LHC Beams at 4* TeV, CERN-ATS-2013-040.

[86] A. Caldwell, D. Kollar, K. Kroninger, *BAT: The Bayesian Analysis Toolkit*, Comput. Phys. Commun. 180 (2009) 2197–2209.

[87] ATLAS Collaboration, *Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s}$ = 13* TeV *with the ATLAS detector*, Phys. Lett. B754 (2016) 302–322.

[88] ATLAS Collaboration, *Studies of Sensitivity to New Dilepton and Ditop Resonances with an Upgraded ATLAS Detector at a High-Luminosity LHC*, ATL-PHYS-PUB-2013-003.